The connection between action and perception

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften

(Dr. rer. nat.)

dem Fachbereich Psychologie der Philipps-Universität Marburg

vorgelegt von

Peter Veto

Geboren in Pécs, Ungarn

28 November 1984

Chemnitz 2018

Vom Fachbereich Psychologie der Philipps-Universität Marburg als Dissertation
am 26.01.2018 angenommen.
Erstgutachterin: Prof. Dr. Anna Schubö
Zweitgutachter: Prof. Dr. Wolfgang Einhäuser
Mündlichen Prüfung: März 2018

Table of contents

Part I.
Cumulus

Introduction and Summary

This thesis consists of three main studies that cover complementary aspects of action-to-perception transfer. In the recent decades, cognitive psychology has started a paradigm shift from its traditional approach to put the stimulus first and treat the action as response to a less one-directional view of perception and action. Quite trivially, action influences perception by changing the external world: we move objects, we locomote or we move our sensory organs. More crucially, action also influences perception internally. Study II and III will address this question directly, by studying perceptual effects of action on physically unchanged stimuli. Study I deals with biological motion. I will argue that the perception of biological motion may present a naturalistic example for direct action-to-perception transfer. The cues of animate locomotion are detected rapidly and effortlessly, and allow quick retrieval of detailed information about the actor, as we related to our immense experience with moving our own bodies in ways that correspond to the physical "laws" which the dynamics of these cues represent. In sum, the studies reported in this thesis provide novel insight on shared action-perception representations, their perceptual consequences and their relation to cognitive models of the world.

In Study I, we showed that biological motion cues distort the perceived size of the actor's figure: a biological motion stimulus is perceived larger than matched control stimuli and lets subsequent stimuli appear smaller. Provided the importance of biological motion, this is in line with other studies that relate subjective importance to perceived size – however, the connection with animate motion has not been reported earlier. If there are shared action-perception representations, do they operate on different representational levels? In study II, we coupled a stimulus that was in competition with another to action more or less strongly. While the degree of action-perception coupling did not affect overt reports of stimulus' visibility, oculomotor measures were modulated. This suggests different degrees of action perception coupling on different representational levels, with varying access to awareness. Does in turn the internal cognitive model of the world penetrate action perception coupling? In study III, we showed that the effect of action-perception congruency on perceptual stability critically depends on the internal cognitive model of action perception coupling. Studies II and III together indicate that no single mechanism or representation can account for all action-perception findings. In the general discussion, I will consider the needed adjustments to current models as well as alternative theoretical approaches.

1. Biological motion perception

For terrestrial animals, the most important feat to recognize is the locomotion of others, the majority of whom are legged creatures. Whether potential prey, predator, mate or a fellow gatherer to share tasks with, recognizing their motion has arguably been crucial for our highly social species throughout its evolutionary history. There are certain levels of this process, starting with recognition, but going far beyond that: from its motion only, we are able to retrieve information about physical as well as personal properties of the actor.

If we consider our normal visual sensory input impoverished (which is to be interpreted in relation to the amount of useful information we are able to retrieve from it), the most common biological motion display will certainly illustrate the logic behind this assumption. Used first by Johansson (1973; 1976), point-light displays depict the movements of the human body's major joints in a highly condensed way. From the 2-dimensional motion of a few dots, we are able to readily recognize complex actions, without any surface information and with no explicitly defined structural connection between the dots.

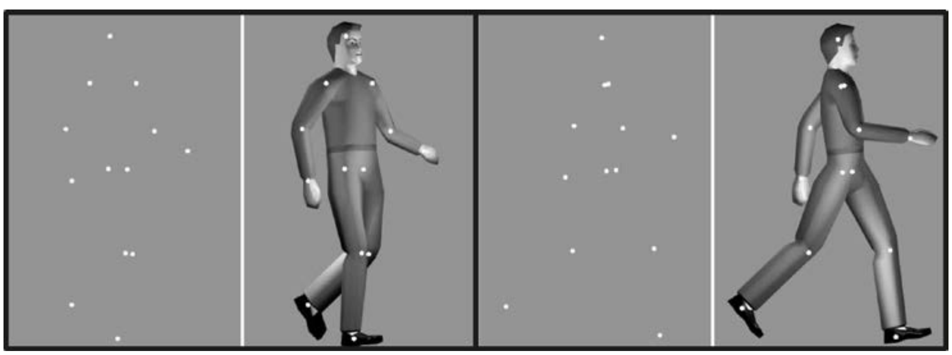

Figure 1. Point-light figures. Note that in an orthographic rendering, there are two possible interpretations due to depth-ambiguity – the illustrations here highlight only one of them. (Text and images have been modified. Source: Vanrie, Dekeyser, & Verfaillie, 2004)

Beyond simple recognition, we are also able to rapidly detect biological motion stimuli, which carries obvious functional significance. As Johansson (1973) anticipated, this ability seems to be "… a highly mechanical, automatic type of visual data treatment." This ability can even be observed when the configural information (structure from motion) that is available in a point-light walker as described above is removed by spatially scrambling the dots, resulting in an isolation of local motion cues. In a search-task, Wang, Zhang, He and Jiang (2010) found a search advantage of such spatially scrambled point-light walkers. That is in line with other studies which demonstrated incidental processing of biological motion stimuli (Thornton & Vuong, 2004) and a rapid modulation of attention by local biological motion cues (Wang, Yang, Shi, & Jiang, 2014) – reinforcing Johansson's original hunch. Point-light walkers do not only draw attention to themselves, but also orient it towards their direction of translation, as demonstrated by Shi, Weng, He, and Jiang (2010). While some studies emphasize the importance of local motion cues in the processing of point-light figures (Chang & Troje, 2009; Mather, Radford, & West, 1992; Saunderes, Suchan, & Troje, 2009; Troje & Westhoff, 2006; Wang *et al.*, 2014), others show that global aspects can be important as well (e.g. Beintema & Lappe, 2002; Bertenthal & Pinto, 1994; Coulson, 2004; Cutting, 1981; Lange & Lappe, 2006; Neri, Morrone, & Burr, 1998; Shiffrar, Lichtey, & Heptulla-Chatterjee, 1997). This makes it clear that both play a role, and probably the most important aspect of the global-local debate is that different cues can contribute in complementary as well as interchangeable manners, suggesting that biological motion should not be treated as a single phenomenon but rather as a combination of many different aspects (Troje, 2008; Troje, 2013).

Indeed, elements to consider seem plentiful. Aside from recognition, rapid detection, and evocation of attentional orienting, biological motion cues are related to a host of detailed information that observers can correctly identify about the actor. Only a couple of examples are the detection of gender from the motion of faces (Hill & Johnston, 2001) and from the motion of other parts of the body (Kozlowski & Cutting, 1977; Mather & Murdoch, 1994; although see also Pollick, Lestou, Ryu, & Cho, 2002), identification of affect (Atkinson, Tunstall, & Dittrich, 2007; Ikeda & Watanabe, 2009; Pollick, Paterson, Bruderlin, & Sanford, 2001) and identity (Loula, Prasad, Harber, & Shiffrar, 2005; Troje, Westhoff, & Lavrov, 2005).

This multiplicity makes it difficult to give an overarching explanation for the phenomenon. Where do these complex skills originate from? One approach is perceptual: since comprehension of the actions of other animals around us is paramount, we evolved to have outstanding visual abilities in this regard, already from an early age (Fox & McDaniel, 1982; Pavlova, Krageloh-Mann, Sokolov, & Birbaumer, 2001), maybe even from birth (Simion, Regolin, & Bulf, 2008; Vallortigara, Regolin, & Marconato, 2005). Some basic cues in biological motion draw attention in an incidental fashion and lead the observer to quickly recognize an animate agent. Then, action-specific details are recognized through a slower perceptual mechanism, where bottom-up processing of distinctive traits and top-down effects of knowledge interact (Troje, 2008; Zacks, 2004) likely in a hierarchical manner (e.g. Ahissar & Hochstein, 2004; Hemeren, 2008).

Another possible explanation might be related to our own experience with similar actions. Note that this explanation is not necessarily at odds with the one outlined in the previous paragraph, but depending on the theoretical viewpoint, may involve crucial differences. Biological movement patterns are largely governed by the principles of biomechanics and physics. For example, the two-thirds power law, first reported for handwriting and drawing movements (Lacquaniti, Terzuolo, & Viviani, 1983), describes that the velocity of movements increases with the radius of curvature, and applies to the motion of walking (Ivanenko, Grasso, Macellari, & Lacquaniti, 2002) as well as to motion perception (Flach, Knoblich, & Prinz, 2004). Also, Fitts's law, which describes a speed/accuracy trade-off in movement, can be observed in perceptual decisions too (Grosjean, Shiffrar, & Knoblich, 2007). Similarly, the perceptual advantage in recognizing one's own movements (e.g. Knoblich & Flach, 2001; Loula *et al.*, 2005) and in recognizing movements that the observer is experienced in executing (Beets, Rösler, & Fiehler, 2010; Casile & Giese, 2006; Hecht, Vogt, & Prinz, 2001) suggests that knowledge about the kinematics of a specific movement can carry over from execution to perception. Most interestingly, the inverse of this statement also appears to be true: a lack of (recent) experience in locomotion impairs its perception (Arrighi, Cartocci, & Burr, 2011).

Practical outlook I – representation of gravity

The importance of gravity in the perception of biological motion displays seems clear from the inversion effect (Sumi, 1984) – when point-light walkers are inverted, the amount of represented visual information remains equal compared to that in upright displays, yet many of the previously listed perceptual abilities related to these figures disappear. The reason might be that the acceleration patterns of the inverted movements (Troje & Westhoff, 2006) represent a situation (negative gravity) that is never observed in reality. Experimental data on imagery of biological motion are scarce (but for some related findings, see Deen & McCarthy, 2010; Grosman & Blake, 2001; Miller & Saygin, 2013), and it would be particularly difficult to design a study where the gravity-related visual cues could be studied separately in imagery. In the perception of the movements of external objects, we seem to incorporate the effects of gravity accurately (e.g. Jörges & López-Moliner, 2017; Lacquaniti, Carrozzo, & Borghese, 1993), suggesting that we might have a fairly precise internal model of it. However, the recent study of Gravano, Zago, and Lacquaniti (2017) shows that in imagery, we do not account for gravity, as if this internal model was not functioning appropriately, or was detached from imagery. This may well explain some of the differences in efficiency of imagery training in various sports (Hall, Rodgers, & Barr, 1990; Sheikh & Korn, 1994), and the accounts of gymnasts where skills with prolonged air-time regularly fail in imagery while they succeed in reality. This is in contrast with the finding that athletes often imagine winning and rarely imagine losing (Hall *et al.*, 1990) – suggesting that there is indeed more to the phenomenon than pessimism. Repeated failures in imagery might in turn affect the confidence of the athlete and studies like that of Gravano and colleagues (2017) might help to understand why this experience is to be expected. Similarly, coaches usually instruct athletes to “use all senses” in imagery practice, while this may not always be appropriate. Mentally practicing a sequence that tends to fail in imagery but succeed in reality might nevertheless be effective in a modality-specific manner: the landing inevitably fails as gravity and the related timing is not incorporated properly in the imagery – but in all other respects, the sequence of the movements can be practiced correctly if the athlete understands that the purpose of this practice is not the timing of the landing but everything else.

Undeniably, the kinematic cues that help us quickly recognize the locomotion of legged animals from their visual appearance are the ones that we have most behavioral experience with. The ballistic acceleration pattern of the feet allows us to identify legged animals of all kinds (Troje & Westhoff, 2006), and we also effectively guess the size of an animal from its stride frequency (Jokisch & Troje, 2003), thanks to the universal physical properties of pendulums under constant gravity. Practicing this very same movement several thousand times a day (Althoff *et al.*, 2017; Bassett *et al.*, 2010) might be an explanation for the generalized filter-like properties of our perceptual "life-detector" (Troje & Westhoff, 2006).

2. Mechanisms of motor and perceptual resonance

The existence of our remarkable perceptual abilities regarding biological motion stimuli does not explain, however, *how* a motor experience transfers to perception. So far we have seen that the actions of animate beings are perceived with high efficiency and I argued that one possible reason behind this is that we have much experience with similar movements. Furthermore, there is ample evidence that such effects are not even restricted to biological types of motion. Action (or planned action, see e.g. Fagioli, Hommel, & Schubotz, 2007) can increase sensitivity to perceptual events that share some features with the action both concurrently (on-line) and with temporal difference (off-line). This has been demonstrated with ambiguous stimuli in the visual (Beets, *et* al., 2010; Mitsumatsu, 2009; Wohlschläger, 2000) and auditory (e.g. Repp & Knoblich, 2007) domains, as well as in imagery (Wexler, Kosslyn, & Berthoz, 1998; Wohlschläger & Wohlschläger, 1998). While binocular rivalry shows some differences from tasks with ambiguous displays, the on-line action-to-perception transfer effect appears similarly in that paradigm, too (Di Pace & Saracini, 2014; Maruya, Yang, & Blake, 2007), or even with unambiguous stimuli that show high perceptual uncertainty (Keetels & Stekelenburg, 2014).

The theories of common coding (Prinz, 1997) and event coding (Müsseler, 1999; Hommel, Müsseler, Aschersleben, & Prinz, 2001) provide a framework for these findings. Since data suggest that there is a generalized connection (with varying specificity, depending on the task and stimulus) between the motor and perceptual domains, these theories imply that both motor and perceptual events are coded in a common representation. This representation only applies to the events' most substantial actuality; their details are coded peripherally, allowing for the varying degree of generalization that we have found in the available experimental evidence. Thus, the theories do not elaborate on the distal part of possible mechanisms (i.e., early sensory processes and late motor processes), but define the connection between domains as a bidirectional flow of information at a central level (theory of event coding), where late stages of perception and early stages of action share a common representational domain (common coding).

So far, our best evidence about the mechanisms that lie behind the information transfer comes from studies on the mirror neuron system (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992) and show similar activation of premotor cortical areas in macaques both when an action is observed and when it is performed. Although more debatable, humans might function similarly (Calvo-Merino, Glaser, Grezes, Passingham, & Haggard, 2005; Calvo-Merino, Grezes, Glaser, Passingham, & Haggard, 2006). These studies led to a wide, though not very accurately defined (Uithol, Rooij, Bekkering, & Haselager, 2011), use of the term “motor resonance”, which entails that observation of actions leads to an internal mirroring that underlies action understanding. This same process, on the other hand, is not unidirectional and so, “perceptual resonance” (the way self-generated action leads to an influence on perception; Schütz-Bosbach & Prinz, 2007) can be viewed in a similar framework.

3. On what level of processing is the connection between action and perception?

While early theories in cognitive psychology treated cognitive processes as linear series of stages (e.g. Sternberg, 1969), where each stage has to end before the next stage would begin, the common coding and event coding approach allows interaction between (partially) linear processes at their highest levels. But where exactly are these levels in a functional sense? The picture we get from the literature is not conclusive. Many studies agree that performed (or imagined) action and the perceptual task need to have a shared dimension for the transfer to take place (e.g. Beets *et* al., 2010; Keetels & Stekelenburg, 2014; Wexler *et* al., 1998; Wohlschläger, 2000; Wohlschläger & Wohlschläger, 1998). In case of studies using mental tasks, there is no kinesthetic information to possibly bias perception; furthermore, due to the necessity of a shared dimension, top-down control likely plays a role. This is interpreted as a process that is similar to attentional capture and named “action capture” by Wohlschläger (2000). However, the shared dimensions in these studies could as well be due to a low-level matching (e.g., the direction or orientation of the events). In the study of Beets and colleagues (2010), participants reported their percept of an ambiguous rotating cylinder by either button presses, or by rotating a manipulandum lever congruently or incongruently to the percept. Results from the manipulandum rotation conditions revealed a congruency effect (stabilization of the percept with congruent motion as compared to incongruent motion), which replicated the findings of Wohlschläger (2000). More interestingly, in an additional condition where percept was reported by button presses while concurrent manual rotation was performed, the action with the manipulandum in a predefined direction did not affect the stability of the percept. This control condition of concurrent, but task-irrelevant, action could give stronger support for the idea that a low-level information flow cannot account for the findings. A possible issue though is that this study used a dual-task for task-irrelevant action conditions only and not for task-relevant action conditions. In the study of Maruya and colleagues (2007), the binocular rivalry paradigm allowed for separately testing effects when the stimulus was dominant vs when it was suppressed from awareness. Their results, showing an effect of action also on the suppressed stimulus, strongly argue for the possibility of a low-level action-to-perception transfer that is outside of top-down control.

So far, no study has shown a clear dissociation between higher and lower relative levels of processing that could unequivocally point to the stage where the shared representations of the common coding theory take place, leaving this part of the model relatively speculative.

4. Experimental studies

All three studies included in this thesis measure perceptual biases, which are caused by action. In all the presented experiments, visual stimulation remains unchanged (or matched in its physical properties, as in Study I) between conditions, supplying a control that unequivocally points to the internal nature of the measured biases.

Here, I introduce the three studies on a basic level, including their results and the immediate aspects of their interpretations. Also included are some considerations, which did not make part of the articles, but may be relevant to the conclusions of the dissertation. Aside from these latter points, the article manuscripts in Part II should be consulted for detailed descriptions. General conclusions regarding what the studies reveal about the inner workings behind the effects are given in the final section of Part I.

Study I (Veto, Einhäuser, & Troje, 2017) demonstrates that an abstract depiction of the most basic form of locomotive action in human life, bipedal walking, is perceived as taking up more physical space than an ecologically invalid depiction of the very same movement. In three experiments, we (i) quantified the phenomenon, (ii) offered a controlled replication to verify that it is indeed the motion and not any static aspect of the stimulus that drives the perceptual bias, and (iii) showed that the perceptual distortion can also be measured indirectly, through a carry-over to simple stimuli.

Study II (Veto, Schütz, & Einhäuser, in press) assesses the effects of the viewer's own movement on the perception of a stimulus that is more or less related to the performed action. A continuous flash-suppression paradigm with eye-tracking allowed us to capture not only the conscious percept of the observers, as expressed in their subjective self-reports, but also the objective measure of their eye-movements which were directly related to the target stimulus. By using this paradigm, we tackled a lower level of action-perception coupling than earlier studies which showed effects of an on-line action-to-perception transfer. The finding that in our paradigm, action only affected eye-movements but not the reported percept, gives novel insight into how action can affect perception on different levels of processing and demonstrates a dissociation between various courses of the transfer.

In Study III (Veto, Uhlig, Troje, & Einhäuser, submitted manuscript), we manipulated our test participants' cognitive models of the coupling between their own actions and the actions' perceivable outcomes. By inducing the assumption of different coupling mechanisms, we measured whether the on-line action-to-perception transfer was merely a result of a direct information flow from one domain to the other or if cognition could penetrate this process. Results showed that the internal model of the viewer plays a significant role in the action-to-perception transfer, which further supports the notion that the transfer does not take place on one specific level, but rather on several different levels of processing.

Study I: Veto, P., Einhäuser, W., & Troje, N. F. (2017). **Biological motion distorts size perception.** *Scientific Reports, 7*(10), 42576; doi: 10.1038/srep42576

Rationale – Study I

Size illusions, where the spatial dimensions of a stimulus are systematically misjudged, can be observed in a wide variety of visual scenarios. Classic examples operate by using simple stimuli to exploit the principles of size constancy and size contrast. In the former, an object of a given retinal size will be perceived as larger, if the assumed viewing distance of the object is greater, as opposed to when it is perceived to be closer to the viewer. Powerful demonstrations of this phenomenon, as e.g. the Ponzo illusion, use stimuli that are of the same size, yet they appear different to us due to the 3rd dimension that is implied in the image. Size contrast, on the other hand, refers to situations where an object appears in the proximity of another, differently sized object. Here, no depth is apparent in the image, but an incidental comparison between the objects leads to the distortion. This comparison can happen simultaneously (as in the Ebbinghaus/Titchener illusion) or with a temporal delay between presentation of the two objects, creating a size adaptation aftereffect (Polsinelli, Milanesi, & Ganesan, 1969).

A less known group of perceptual distortions in size judgments, however, operate on higher levels of perception, using object properties that are of internal nature: in sum, an object that is of special interest to the viewer tends to look larger, and there is a (sometimes bidirectional) relation between perceived size and subjective value (e.g. Blaker & van Vugt, 2014; Dubois, Rucker, & Galinsky, 2011; Duguid & Goncalo, 2012; Marsh, Yu, Schechter, & Blair, 2009; Masters, Poolton, & van der Kamp, 2010; Meier, Robinson, & Caven, 2008; Murray & Schmitz, 2011; Silvera, Josephs, & Giesler, 2002; Yap, Mason, & Ames, 2013; Veltkamp, Aarts, & Custers, 2008). These phenomena would fit the general definition of illusions; however, due to reasons explained below, they are usually not referred to as such. Although several studies have shown effects of this category, they are typically also not even mentioned along with the previous examples that were based on size contrast and constancy. This may be due to their subliminal character: instead of a simple demonstration, this category of perceptual distortions requires some kind of measurement to become obvious. Furthermore – and theoretically more interestingly – while the former group of illusions relies on visual cues that are part of the visual scene and, under normal circumstances, help us make better perceptual judgments, the latter group of perceptual distortions originate from some *internalized* property of the stimulus.

Biological motion stimuli are universally important, as shown by that they are perceived rapidly (Johansson, 1976; Jokisch, Daum, Suchan, & Troje, 2005; Wang *et al.*, 2014) and incidentally (Thornton & Vuong, 2004; Veto, Thill, & Hemeren, 2013), we have an innate sensitivity to perceiving biological motion (Simion *et al.*, 2008; Vallortigara *et al.*, 2005), and human observers are able to retrieve nuanced details about the performer of the motion (Barclay, Cutting, & Kozlowski, 1978; Montepare, Goldstein, & Clausen, 1987; Troje *et al.*, 2005). Given the above, our hypothesis logically follows: if important stimuli tend to seem larger to the observer and biological motion stimuli inherently enjoy preferential processing in the visual system, then biological motion stimuli are also likely to appear larger than a matched control stimulus. We investigated this question in three experiments.

Paradigm – Study I

First, we tested sixteen participants in a paradigm, where for each trial they briefly viewed a target figure followed by a mask to prevent reliance on afterimages. Then, in an adjustment task, participants indicated the size of the previously seen figure both in width and in height. The target figure varied in size (but kept its natural proportions) and consisted of a point-light walker from a frontal view, with either upright or inverted orientation. We found that the upright displays were perceived as significantly larger than the inverted, but otherwise identical, control stimuli.

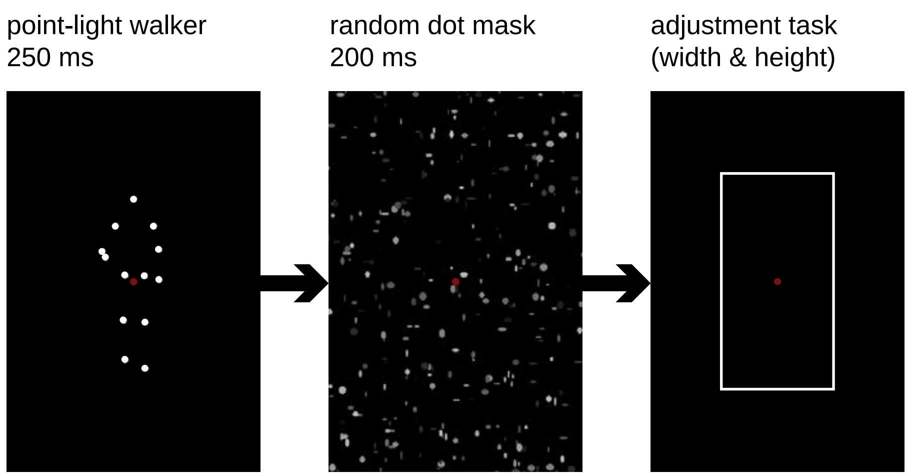

Figure 2. Paradigm, Study I, Experiment 1. Sequence of a single trial (here with upright point-light walker).

Point-light figures present motion stimuli by eliminating the effects of surface cues and also reducing the amount of configural information in the display. Local motion cues, particularly of the limbs, play a crucial role in detecting biological motion (Hirai, Chang, Saunders, & Troje, 2011; Troje & Chang, 2013; Troje & Westhoff, 2006). Since acceleration – driven by gravity – is of major importance in the detection of these cues (Chang & Troje, 2009), inverted figures present an ecologically impossible and thus invalid constellation. Furthermore, the global configuration of the walker also exhibits an inversion effect (Troje & Westhoff, 2006). Consequently, inverting the target stimulus serves as ideal control, where all physical properties of the stimulus remain equal, but the motion cues lose their ecological relevance and do not elicit the same kind of preferential visual processing as upright walkers do. However, due to the inversion effect on the global configuration, our results so far left the question open, whether the measured distortion in perceived size was due to the biological motion cues in the display, or the configurational difference between upright and inverted figures.

In a second experiment, we therefore tested twenty-four participants in a similar paradigm, where the original stimuli were extended by the additional condition of static displays (with both upright and inverted orientations). Here, we found no difference in size judgments on static figures, while the dynamic trials replicated the findings of our first experiment.

In a third experiment, we eliminated the possibility of an unknown response bias that could affect upright walkers differently than inverted ones. If the upright displays appear larger indeed, then subsequent stimuli should be perceived as smaller due to a size contrast effect, as described earlier. Participants viewed a pair of point-light walkers (one of them always upright, the other inverted), followed by the target stimuli (a pair of simple discs; see Figure 3). While the point-light figures were fixed in size, the relative diameters of the two targets varied in five conditions. Participants were explicitly instructed to ignore the point-light figures and only focus on the task, in which they had to pick the larger of the two discs. Results showed that target discs were indeed judged less frequently as the larger of the pair, when they were preceded by the upright walker.

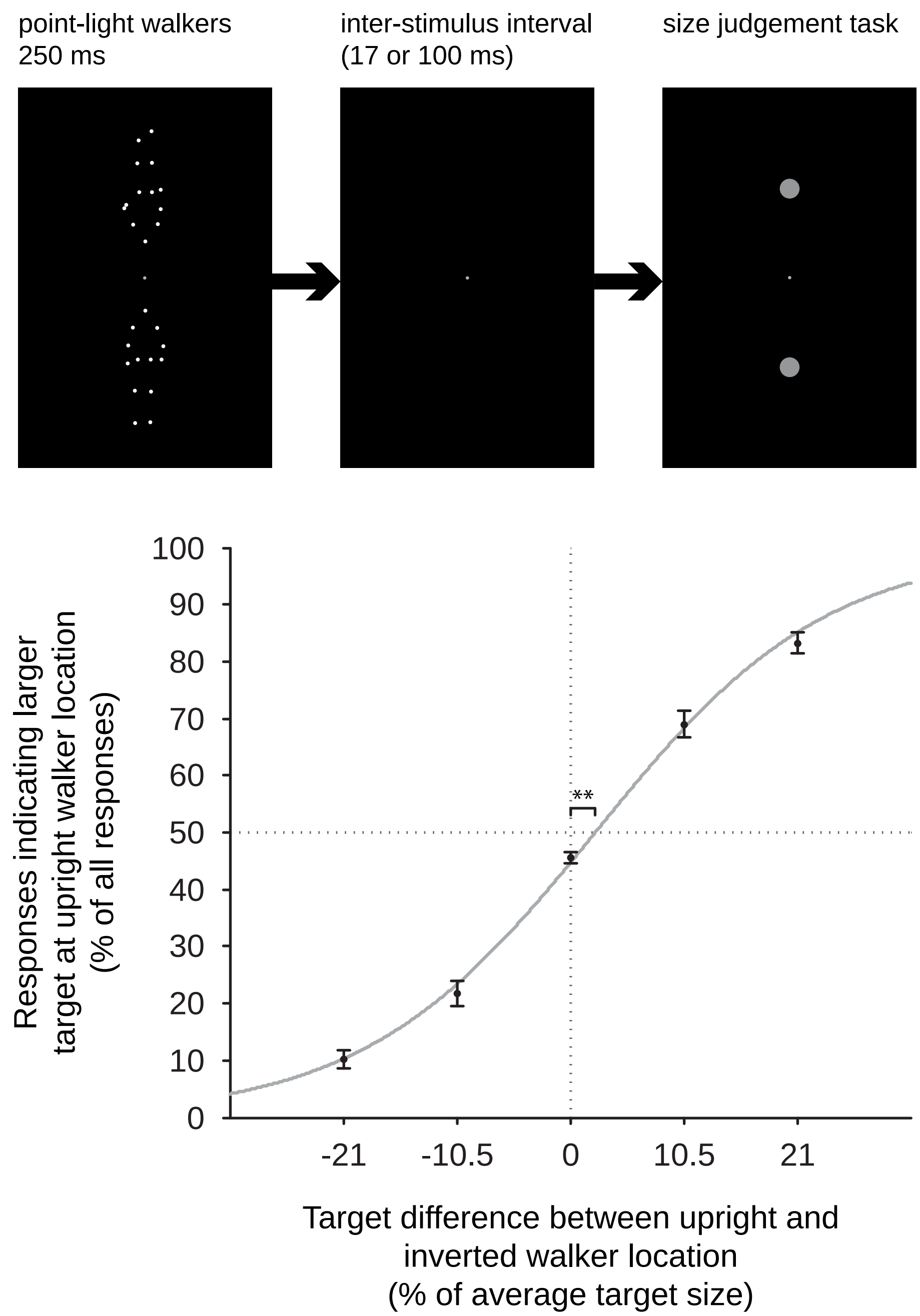

Figure 3. Paradigm and results, Study I, Experiment 3. Paradigm (*top*): Sequence of a single trial (here with upright point-light walker in the lower position and inverted point-light walker in the upper position. Results (*bottom*): Percent of responses indicating that the target preceded by an upright walker was larger plotted against the difference between target (disc) sizes. Means per condition with fitted psychometric function. Error bars show s.e.m. Asterisks indicate significant difference of point of subjective equality at $p < 0.01$.

General results – Study I

The three experiments together provide evidence of a distortion in size perception caused by biological motion cues. This is in line with the literature that shows how biological motion stimuli are processed preferentially by the visual system at an early stage (Jokisch *et al.*, 2005; Wang *et al.*, 2014), eliciting incidental processing (Thornton & Vuong, 2004) and reflexive attentional orienting (Shi *et al.*, 2010). Since our experiments were the first of their kind, some possible vulnerabilities of the interpretation still need to be mentioned.

To avoid a possible effect of hemispatial asymmetries due to an interaction between the attentional orienting response and local vs global processing (Van Vleet, Hoang-duc, DeGutis, & Robertson, 2011) we used a laterally symmetric design, both with regards to the walkers (facing towards or away from the viewer, instead being viewed from the side) as well as to the relationship between the displays (see Figures 2 & 3). This way, however, our results might be connected to the facing-the-viewer bias (Vanrie, Dekeyser, & Verfaillie, 2004), which posits that the depth ambiguity of point-light walkers is more likely to be resolved with the interpretation of the walker facing the viewer and translating towards him or her than with the interpretation of the walker facing away from the viewer. Our experiments cannot answer whether such a connection exists – nevertheless, if it does, the causal direction between the two phenomena would also be unclear.

Results of the first experiment might as well be explained by assuming that the contours of the body are ‘filled-in’ for the upright walkers, but not for the inverted ones, which would also make the upright figures appear larger. Based on this reasoning, on the other hand, results of the second experiment would be difficult to interpret, since the human shape is clearly recognizable in the static stimuli as well. Further experimentation with spatially scrambled walkers or with the isolated motion of the limbs could possibly resolve this issue.

Finally, one point that we did not yet consider at the time of choosing our stimulus, is the possible role of sex differences. Our stimulus was based on the movements of a male actor (Vanrie & Verfaillie, 2004), while our participant population was dominantly females. Women might be more sensitive to perceiving biological motion stimuli (Anderson *et al.*, 2013), and masculine traits and displayed social power could also drive an increase in perceived size (e.g. Murray & Schmitz, 2006; Blaker & van Vugt, 2014; Yap *et al.*, 2013; Marsch, Yu, Schechter, & Blair, 2009). If this connection would account for our findings, a similar experiment using a feminine stimulus should result in an opposite pattern.

As long as the above considerations remain unanswered, our best interpretation of the results is that the ecological importance of biological motion stimuli creates a positive distortion in its perceived size.

Study II: Veto, P., Schütz, I., & Einhäuser, W. (in press). **Continuous flash suppression: Manual action affects eye movements but not the reported percept.** *Journal of Vision*

Rationale – Study II

So far we have seen that the ecological importance of an action can change the way the action is perceived. Perception, however, is also influenced by the own actions of the perceiver, particularly when those actions are related to an observed motion. Such effects have been demonstrated with actions that precede a visual task (off-line effects, see Schütz-Bosbach & Prinz, 2007), or with concurrent actions (on-line effects). Studies of this latter group have indicated that performed action and observed movement need to share some attributes, or else perception would not be biased by performing an action at the same time. For example, studies found effects of manual rotation on mental rotation only when the two happened along an axis of the same orientation (Wohlschläger & Wohlschläger, 1998), or occurred in the same direction (Wexler *et al.*, 1998). Perception of external events (instead of an imaginary visual task) also led to similar findings (see Wohlschläger, 2000; Beets *et al.*, 2010; Keetels & Stekelenburg, 2014). To our knowledge though, no study so far has measured the effect of the type or degree of coupling between action and perception on the action-to-perception transfer. For example, experiments using ambiguous stimuli have either coupled the movement dynamics of a stimulus to test participants' own movements (e.g. Mitsumatsu, 2009), or required that participants report on their percept by means of a motor task (as, among others, in the experiment of Beets and colleagues, 2010), where the dynamics of the response was independent of stimulus dynamics. Since synchronicity seems crucial in our perceptual understanding of the relationships between different events in the external world as well as between our own actions and their causes or consequences (e.g. Aschersleben & Prinz, 1995; Stephen, Stepp, Dixon, & Turvey, 2008), we created a paradigm where an on-line action-to-perception transfer is to be expected, and varied the degree of coupling between action and stimulus movement. Using an ambiguous stimulus and testing whether the action-to-perception transfer depends on the dynamic relationship between modalities is only possible if the action can be connected to one perceptual interpretation of the stimulus independent of the other, competing, percept. For example, applying two distinct gratings with opposite motion directions in a classic binocular-rivalry paradigm would be problematic, as coupling an action to the motion of one grating would likely be interpreted as a coupling to the other.

Paradigm – Study II

To this end, we used a continuous flash suppression paradigm, where a drifting grating stimulus was presented to one eye and a salient Mondrian suppressor to the other. In this setup, the inter-ocular conflict is highly biased due to the low-level saliency of the continuously changing suppressor (novel images presented at 10 Hz), and the target stimulus can only be perceived when endogenous attention is deployed. Test participants pressed and held a button every time the target (grating) stimulus gained perceptual dominance, while in three conditions they (i) rotated a manipulandum device that governed the motion dynamics of the grating in a direct manner (*coupled action* condition), (ii) performed the same action, while the grating was only loosely coupled to their hand movement (*decoupled action* condition), or (iii) performed no action, while the grating's translation followed a similar dynamics as in the other conditions (*no action* condition).

Aside from participants' subjective self-report, we also measured their eye-movements: a horizontally translating grating elicits an optokinetic nystagmus (OKN) response with varying relation to the stimulus velocity (gain). Although the OKN is a motor response per se, these eye-movements only appear when the grating is visible; under complete suppression, eye-movement behavior is dominated by fixations and their connecting saccades. This way, we gained both a subjective and an objective measure of participants' perception of the target stimulus. Our hypothesis was that due to the action-to-perception transfer, the overall perceived strength of the grating stimulus would increase when manual action is performed as opposed to when no action is required. This would be expressed in an increase in the total duration of participants' button presses as well as in an increase in the overall OKN gain. As for the effects of the degree of coupling between action and stimulus, the relationship between the decoupled action condition and the other two conditions would show, whether (a) the degree of coupling has an absolute importance where only directly coupled action leads to a significant transfer effect, (b) the degree of coupling has no effect and the transfer effect can be equally observed with either coupled or decoupled action, or (c) the degree of coupling has a gradual effect on the action-to-perception transfer, and the increase in perceptual strength in the decoupled action condition is between that of the coupled action and no action conditions.

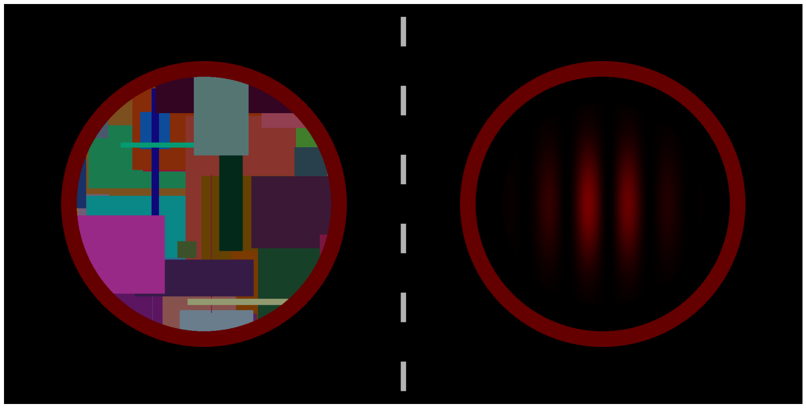

Figure 4. Stimuli, Study II. Stimulus display. The dashed line was added to indicate the separation between the two eyes' stimuli.

Results – Study II

We found no effect of action on the reported perception of the grating, while eye-movements revealed a significant effect, with a linear correspondence to the degree of coupling between action and stimulus dynamics. The implications of these findings are manifold.

First, there is a clear conflict between the numerous earlier studies to show an action-to-perception transfer (as measured by the reported percept) in highly similar tasks and the lack of such an effect in our study. This difference is problematic to dismiss by assuming a lack of power in our case. First, button-press results do not show any marginal difference between conditions – that case could have pointed to the expected pattern simply lacking statistical significance. Second, in a comparable paradigm, our study tested 24 participants while in the experiments of Wohlschläger (2000; N = 5), Maruya, Yang, and Blake (2007; N = 5) or Beets and colleagues (2010; N = 11) smaller sample sizes yielded a clear effect. Third, the eye-movement measurements of our study revealed a significant difference. These reasons together make it highly unlikely that the lack of an action-to-perception transfer in the conscious percept in this experiment was an accidental result. Instead, the resolution might be found in the – so far largely unexplored – differences in how our paradigm affects selective attention as opposed to all earlier studies. Selective attentional control is weaker in binocular rivalry than in the perception of ambiguous figures (Meng & Tong, 2006) and while continuous flash suppression is a case of interocular competition too, it also shows differences that go beyond being merely a stronger case of biased binocular rivalry (see Tsuchiya, Koch, Gilroy, & Blake, 2006). Although our experiment does not directly test such differences and thus, any detailed conclusion in this regard would be unsubstantiated, our findings on the action-to-perception transfer show that the conscious percept is affected differently in continuous flash suppression than in the rivalry paradigms used in earlier studies.

Second, the separation between effects on the conscious percept and on eye-movements alludes to a dissociation between how higher and lower levels of processing are affected by the action-to-perception transfer. This could either mean that different mechanisms are behind the two effects or that the same mechanism is affected at different sensitivities.

Finally, if the reason behind the action-to-perception transfer is the existence of shared representations between the motor and perceptual domains (Prinz, 1997), then the dissociation in our results can be interpreted as a consequence of the existence of several such representations on different levels.

Study III: Veto, P., Uhlig, M., Troje, N. F., & Einhäuser, W. (submitted manuscript). **What you see is what you expect: Cognitive assumptions influence the action-to-perception transfer in ambiguous perception.**

Rationale – Study III

In everyday life, we often use tools without much thought about how they actually work in order to achieve the goal we use them for. Initially though, some kind of understanding has to take place to allow expectations to form. One way for this to happen is through trial-and-error exploration of the input and output of the device (e.g., "which of the two entangled cords do I have to pull to close the window blinds?").

Alternatively, knowledge of the coupling between input and output can also come through learning about the internal workings of the tool. This is most prominent when the connection is not self-explanatory at first sight. While understanding the steering mechanics of a bicycle seems obvious to most adults, the controls of an airplane may take some learning to master. The handlebars of the bicycle and the rudder pedals of the (most simplistic) airplane work in similar ways, but the coupling between the operator's input and the outcome is reversed. In case of the airplane, the reverse coupling might be easier to grasp when the connecting wire between the pedals and the ipsilateral side of the rudder is kept in mind (Figure 5). With many other common tools too, the connection between action and its outcome is arbitrary, e.g. in how the spatial configuration of a rack and pinion defines which directional translation a given rotation is mapped to.

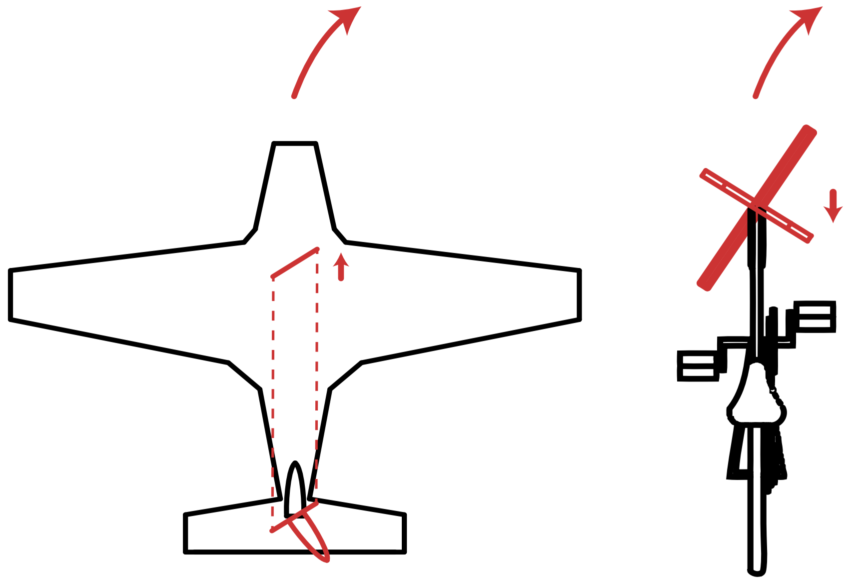

Figure 5. Illustration of simple mechanical systems. Although the coupling is arbitrary and can change from one system to the other, we are able to learn them at a level where finding the right direction already at the first movement does not require cognitive effort. However, the learning can be facilitated by knowledge of the mechanic layout.

As we have seen earlier, many experimental findings demonstrated that action can affect perception and such effects take place with higher likelihood when the perceived motion is related to the action. In examples like that of the previous paragraph, action and its perceived outcome are obviously coupled, but their mappings are complex, at times even conflicting. So what do we rely on, when we use tools correctly, based on decisive predictions? If the shared representations between action and perception are on a low level, where top-down effects have no influence, then these predictions must come from straightforward contingencies. Indeed, results from Maruya and colleagues (2007) suggest that a direct information flow between the modalities is possible, as they found an action-to-perception transfer effect even when the action-coupled stimulus was outside of awareness. Similarly, Study II demonstrated that in a situation where the conscious percept was unaffected by coupled action, eye-movements still displayed an effect. Other studies, like those of Wohlschläger (2000) or Beets and colleagues (2010), point in the opposite direction and suggest that the transfer should only take place when stimulus and action share a cognitively defined dimension, thus requiring cognitive mediation. In the present study, we aim to dissolve this contradiction by a paradigm that tests participants' cognitive model of action-perception coupling without biasing the immediate contingency between stimulus and performed action.

Paradigm – Study III

The *internal model* of coupling was induced by a rendered depiction of a 3 dimensional rotating cylinder, connected to a circulating lever (that resembled the manipulandum handle used by participants in the subsequent perceptual task). The connection was either through a belt-drive mechanism (“belt” condition; the rotation direction of cylinder and lever were the same) or through cogwheels (“gear” condition; opposite rotation directions). After participants had studied the mechanical model of the assigned condition, they completed four test-blocks with an ambiguous version of the cylinder (Figure 6b). Their task was to report their percept of the cylinder, by rotating the manipulandum lever in the same or opposite direction (orders counterbalanced) as the lever of the mechanical model would rotate (red bar of Figure 6a). This way, the *match* between stimulus and action direction was either congruent (same direction instruction in the belt and opposite direction instruction in the gear condition) or incongruent (opposite direction instruction in the belt and same direction instruction in the gear condition) with the perceived cylinder rotation (Figure 6d). The same procedure was then repeated with the other mechanical model (order counterbalanced between participants; for details, see Table 2 in the article). Perceptual stability was assessed based on the lengths of intervals when the reported direction remained unchanged.

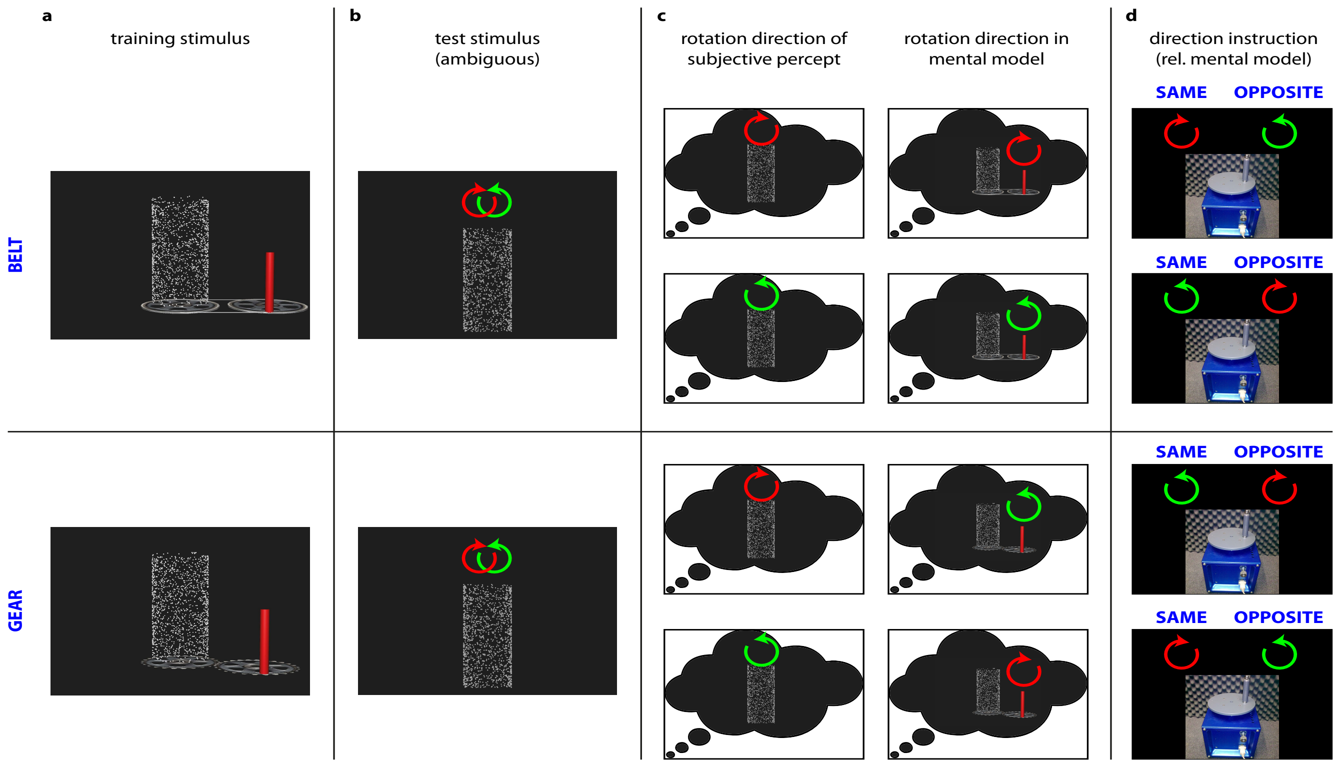

Figure 6. Stimuli, percept, and task, Study III. a) First, in separate blocks (blocks 1 & 6, see table 2), participants were introduced to the mechanical model ("belt" or "gear" layout). For 30 seconds, they controlled the displayed motion with the manipulandum. Then, 20 seconds of unambiguous motion followed (the cylinder and mechanical model rotated with occasional switches in direction), where observers had to report the rotation of the red handle in accordance with the subsequent experimental block ("same direction instruction" or "opposite direction instruction"). For the last 20 seconds of training, the red handle bar disappeared and the mechanics was covered by a virtual desk, while the task remained unchanged. **b)** All test blocks showed the same, ambiguous, motion cylinder for 3 minutes each. **c)** Two possible perceptual interpretations of the test stimulus (clockwise and counter-clockwise). Participants had to respond to the imagined motion of the red lever, as it related to their current percept. **d)** Instruction (manipulandum rotation in the same or opposite direction as that of the red lever in the mental model). Note that in the "belt" condition, the same/opposite direction instruction leads to congruency/incongruency between perceived and performed rotation, while this relationship is reversed in the "gear" condition.

If the action-to-perception transfer is dominated by perception (direct information flow between domains), then regardless of the induced *internal model*, the effect of *match* between perceived stimulus direction and performed action will lead to the same congruency effect as found by Beets and colleagues (2010). However, if cognition plays a major role in how the transfer is formed, results from the two models should diverge: the increase in perceptual stability in congruent directions should reflect the instruction on the mechanical model's lever and not the actually perceived direction. This would lead to a reversed pattern in the gear condition and an interaction between factors *internal model* and *match* in our design (see Figure 7).

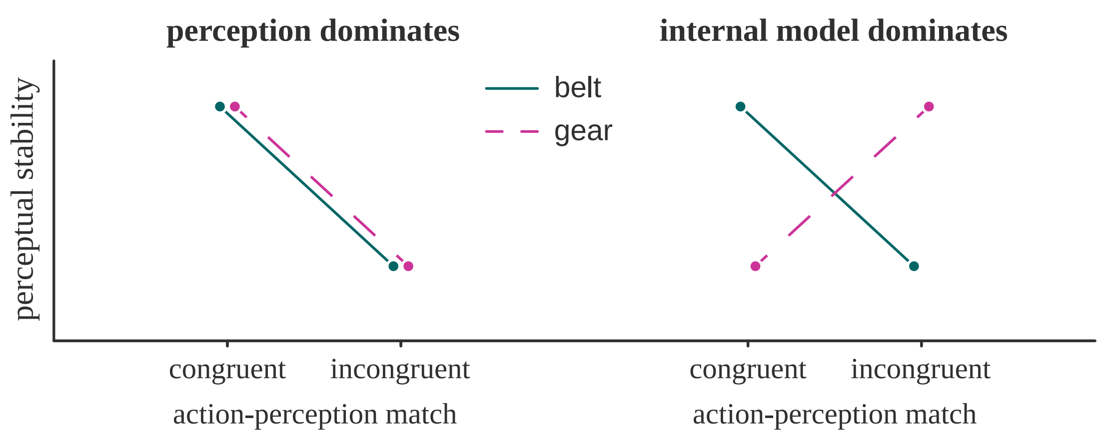

Figure 7. Hypotheses, Study III. Expected effects of *internal model* and *match* (between percept and action) on perceptual stability, if action-perception coupling is not under cognitive influence (*left*) or dominated by the cognitive model (*right*).

Results – Study III

A significant interaction between the two factors showed an influence of the cognitive model of coupling on the action-to-perception transfer. The reversal of the congruency effect was not complete however: the match between perceived direction and performed action only showed a significant difference between congruent and incongruent directions in the belt, but not in the gear condition (Figure 8). These results together indicate that while cognition significantly influences the action-to-perception transfer, it is not the sole source of the effect. The lack of a significant effect in the gear condition alludes to a cancellation of the natural action-perception congruency effect by an equally large influence of the assumed mechanical model.

In line with the conflicting results in the literature and with the findings of Study II, we conclude that the observed action-to-perception transfer phenomenon cannot be explained by a model that assumes a single source of the effect. Study III also shows that when using simple tools, we do not only rely on direct sensorimotor contingencies, but may also incorporate internal assumptions that exert top-down influence on our perceptual expectations.

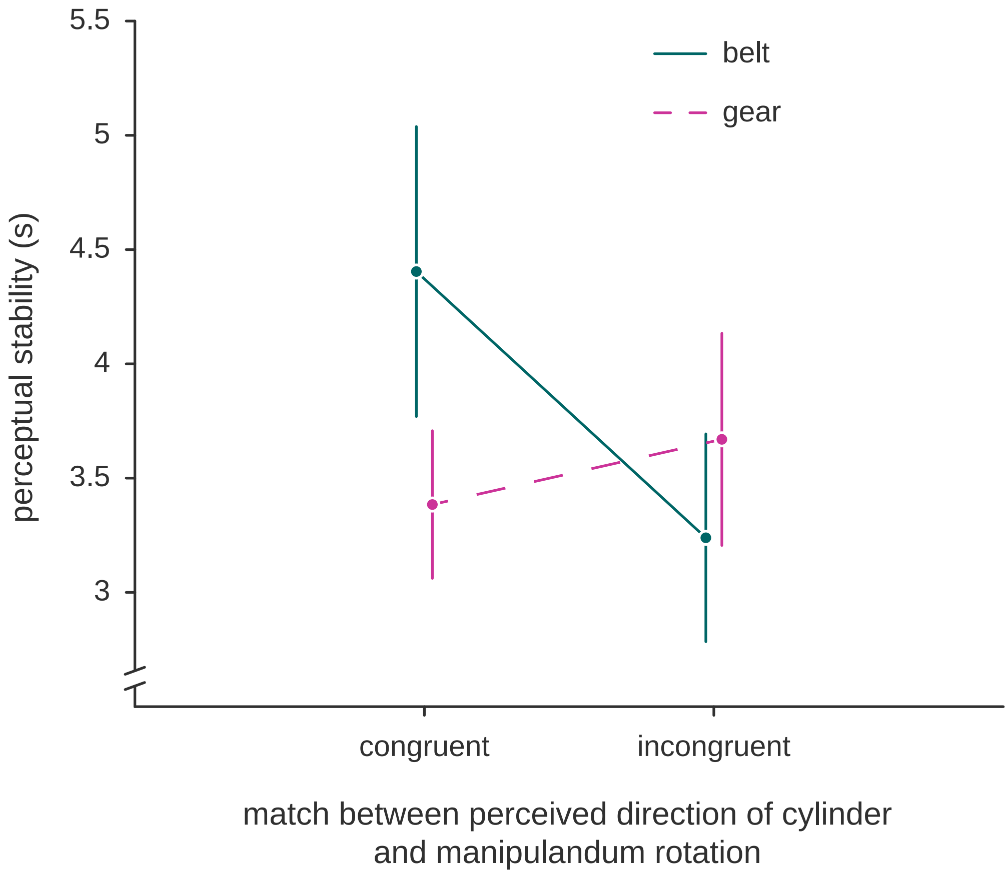

Figure 8. Results, Study III. Perceptual stability for each condition, averaged across participants. Error bars show standard errors of the mean.

Practical outlook II – perceptual expectations in real life

It appears so that using tools relies on learned contingencies between action and its outcome, as well as on an internal model. These two sources can interfere positively or negatively and in case of a positive interference, a stronger expectation is formed. Using a simple tool involves such expectations, where repeated practice leads to a solid internal model, which in turn increases efficiency by freeing up resources. For example, a child learning to use a screw driver might initially have to use trial-and-error to find the correct direction, until the mechanism is learned.

The skill of driving a car is mastered similarly, where up to the 1990s, practically all processes were highly predictable under normal circumstances. Since then, more and more controls are taking effect through a computerized mediation, leading to less predictability between action and perceivable outcome. Although the computer itself may be highly predictable, the user who does not understand its workings would not be able to anticipate its reactions. The most simple example is how a hybrid powertrain reacts to a given amount of accelerator input differently depending on the state of the system, e.g. whether the internal combustion engine is currently on or off. With the increase of computerization from traction control through driver assistance systems to conditional automation, the uneducated driver is facing increasing unpredictability. While initial stages of this trend might only keep enthusiastic drivers away and do not affect safety negatively in those who adapt to the new systems, later stages have revealed two sources of issues, both of which can be traced back to our logic regarding sensorimotor contingencies. First, drivers experiencing a decrease or complete loss in coupling between their actions and the actions' effects might lose trust in the machine (Abraham *et al.*, 2017) and end up not using these systems (Kidd, Cicchino, Reagan, & Kerfoot, 2017). Second, and probably more importantly, the lack of connection can lead to a decrease in attention and a drastic drop of involvement (Geitner *et al.*, 2017; Reimer *et al.*, 2016), where unreasonable trust is given to a personified machine (Inagaki & Itoh, 2013; Waytz, Heafner, & Epley, 2014). Study III suggests that increasing the knowledge about how the machine operates (building the "internal model of coupling") could help with these issues, which is in line with the findings of studies in the applied field (Sonoda & Wada, 2017; Thill, Hemeren, & Nilsson, 2014).

5. General discussion

Study I answers a simple question, which, to my knowledge, has not been asked before. Regardless of whether the reason behind our remarkable sensitivity to perceiving the movements of animate entities is purely perceptual in nature or comes from the internalized effects of motor practice in the perceiver, our widely demonstrated skills in biological motion perception allude to a high ecological importance of its visual cues. As studies with other types of stimuli have shown, important objects tend to look larger to the observer, and now we have evidence that biological motion displays are no exception. Although in the introduction, I argued for the view of considering action-influence as a significant factor in the forming of the preferential processing of biological motion, the line of thought as well as the empirical evidence presented there only attempt to explain the perceptual abilities around the recognition and identification of animate entities. The reason behind the size-distortion effect remains just as mysterious as in the earlier studies using different stimuli. Arguments can be made for a relationship with an internal representational space, as in the "SNARC" effect and its relation to the vertical dimension (Dehaene, Bossini, & Giraux, 1993; Ito & Hatta, 2004), but this connection seems more or less plausible only for social stimuli that operate with power and vertical hierarchy. If a similar affinity in the use of a male walker and predominantly female observers in Study I is a contributor to the results, then the argument might also apply to the case of biological motion perception. Having no other speculative alternative, we can assume that with some likelihood, the explanation may lie in that all evaluations of subjective importance are organized in a mental space that interacts with the representation of physical space.

Studies II and III test the properties of the on-line action-to-perception transfer. The theories of common coding (Prinz, 1997) and event coding (Hommel *et al.*, 2001) suggest that the effect takes place on higher levels of processing and involves only the general properties of an event, while the details are coded distally and do not interact between domains. While results of many studies support these ideas, some raise questions about the level of processing where the transfer might take place and about the more or less cognitive nature of the transfer (see Sections 2 & 3). The original aim of Study II was to assess the contribution of the degree of coupling between action and perception on the action-to-perception transfer. To achieve this, we implemented our conditions of variable coupling in a continuous flash suppression paradigm that has not been used in a similar context so far. This choice was primarily due to the practical considerations of connecting movement to only one state of a bistable percept. Incidentally however, it also led to further and unexpected findings. First, seeing a gradual effect of the degree of coupling on the OKN gain suggests that in our paradigm, the action-to-perception transfer did not only include the general presence and direction of rotation, but also the details of its dynamics. Second, and yet more contrary to the predictions of the common coding principle, we found no effects on the experiment's higher-level measure of percept (the conscious self-report), possibly due to the selective attentional involvement in the CFS paradigm. This alone would talk more to the nature of CFS than support or reject the common coding model; however, simultaneously finding significant effects in a lower-level measure means that the assumptions about the hierarchical nature of the transfer may need revisiting. In line with the findings of Maruya and colleagues (2007), our results also show that a transfer effect outside of awareness seems possible and that top-down control is not necessary.

Wohlschläger (2000) argues, also based on experimental evidence, for "action capture", where a low-level, direct, information flow is unlikely to account for the effects. Results of Study III are partly in support of his views, as they demonstrate that a cognitive model of the effects of action can change the perceptual bias of action, even with a constant visual stimulus and no explicit exposure to the sensorimotor contingency that could explain the results on a lower level. However, the pattern was not completely reversed by the internal model in Study III, showing that the lower-level coupling between concurrent action and perception can also have an effect through a separate mechanism. This, taken together with the findings of Study II, paints a picture that is less in favor of the original arguments of Wohlschläger (2000).

It seems as if the more data we have, the tougher it becomes to locate the action-to-perception transfer in processing hierarchy. With each new paradigm, the question is asked slightly differently and the answers keep pushing the boundaries of the common coding theory as they were originally described. One possible resolution is to assume that several levels of processing can create connections between action and perception and the model should be flexible in this regard. The common coding theory, in its original form, cannot account for all the experimental evidence that has gathered in the two decades since its creation. Modifying it by positing that the shared representations do not necessarily have to be at the proximal end of the two processing streams but could occur on many levels (even independently), would give the ability of explaining the otherwise contradicting results of this thesis as well as of earlier studies. This also fits the general trend in cognitive science from rigid models to more flexible ones. On the other hand, it would create the necessity of a host of "new" representations, possibly one for each new way of testing – and if that was indeed the case, it would challenge the very purpose of modeling. Furthermore, such a liberation of the common coding theory would evidently bring up a question, which has in fact been lingering around since the earliest days of the model. While the original question was along the lines of "on what level does the transfer take place?", now we should ask "what is the lowest level, where the transfer might occur?"

So far, I have considered all theoretical questions of this text in a representational framework. Beyond being the most common in today's psychological approaches, this viewpoint seems to formulate the questions of my experiments well. However, discarding the assumption that any experience is based on an internal representation could eliminate the issues detailed above (and possibly raise other ones). While the related debates around the existence or nonexistence of a "Cartesian theater" reach far back in the history of philosophy, when it comes to psychology, the non-representational viewpoint is often traced back to William James (1912). Oddly enough, the representational framework also finds a starting point in his writings (James, 1890), at least for psychologists. One may get a clearer and more contemporary picture of what this direction of thought entails from the writings of Gibson (2015), where he argues that perception is the starting point that we need to understand first, keeping in mind that organisms are mobile and perception operates in service of action. In this functional account of perception the two domains are not even really separate from each other, and the perception of invariant structures (objects) requires motion through time. Due to this, taking the retinal image at a given time point as the basis of visual perception is incorrect and misses the most important, dynamic and relational, pieces of information about the environment. Furthermore, if this information is in the interaction with the environment, there is no need for complex internal computations to model the world from an impoverished input. This way, all my earlier contemplations in this script about how our participants' subjective percepts are biased are misguided, as the individual's percept is not subjective but simply relational to his or her self, and these relations are not inferred but perceived directly. If, for example, the stimulus resembles a person, the relationship (or "affordance") will be very different than in the case of meaningless dots.

This view aims to completely get rid of the issues with representations by positing that perception is an interactive process where the external world is used as its own representation, through constant interaction with it. This may as well have interesting connections to some puzzling findings, described in "Practical outlook I – representation of gravity" (p 8). Gravity is by far the most constant invariant in our daily experience, so we can always use it accurately in real-world tasks: in any ballistic movement we can detect the cues of gravity and use them for accurate predictions. However, in imagery, these initial cues are missing and the imagined outcomes will be highly unrealistic. Would this also mean though that everything else (which we can imagine well) *has* internal representations – suggesting a hybrid model that is not radical in either direction (see below; O'Regan & Noë, 2001)?

While leaving open a lot of questions as for *how* direct perception actually happens, there are many practical implications of this model, which might be the reason for the current rise in popularity of similar ideas where actions is inseparable from perception and cognition (Engel, Maye, Kurthen, & König, 2013).

Note also a common confusion of terms, due to them being vaguely defined. In psychology, we often contrast behavioral and neural measures, which serves practical purposes when describing methods. But in essence, what is *not* a behavior? In biology, the term is more clearly defined: on the level of a retinal receptor cell, for example, activities induced by light are the cell's behavior. The lowest possible levels of sensation are also behavioral.

Using a somewhat similar (albeit far wider reaching) logic, O'Regan and Noë (2001) propose a framework where "seeing is a way of acting". The crucial term in their theory is sensorimotor contingency: instead of representations based on sensation, they claim that all experience comes from an exploratory activity, where sensorimotor contingencies contain the knowledge that mediates in this exploration. The argument is that from the perspective of the brain, there is no difference from neural input from one source or another. All differentiation comes from previous experience of the structure of how motor actions induced (or rather, co-occurred with) changes in the given sensory input; in other words, from the sensorimotor contingency.

This view offers remarkable flexibility. For example, no complicated mechanism is needed to explain the differences between senses. The physical relationship between action and sensory input is itself very different between the modalities, and the same mechanism that extracts these contingencies would uncover the regular relationships (or "laws", as the authors call them) in both cases, despite the differences. In this framework, the interaction between action and perception happens on all levels at all times and do not require common coding or representations of any sort – seemingly solving the issues that we considered previously. Results of Study III show effects of cognition without any direct sensorimotor coupling behind – so how would these results fit the model? According to the authors, the knowledge of sensorimotor contingencies can extend to arbitrary levels of abstraction. This does not mean that there is an actual, pictorial, model somewhere in the mind, or following the authors' example "… as though, in order to generate letters on one's screen, the computer had to have little letters floating around in its electronics somewhere" (O'Regan & Noë, 2001). Instead, it means that the possible effects of manipulations are considered based on the extraction of "laws" (not necessarily the direct effect itself) from previous experience. This theory somehow bridges the nonrepresentational and representational frameworks at this point. Conceivably the most authentic definition of radical empiricism is in James' *Essays in Radical Empiricism* (1912): "To be radical, an empiricism must neither admit into its constructions any element that is not directly experienced, nor exclude from them any element that is directly experienced." Taking this definition as a starting point, one might argue both for and against the inclusion of sensorimotor contingencies – or even more ambiguously, of the knowledge of them – into the category of directly experienced elements, based on the detailed interpretations of these terms.

Ultimately, these questions have crucial importance, as they deliver the most pervasive impact on our personal 'theories of everything'. For the empirical psychologist however, this level of philosophical abstraction might seem as a mere collection of terminologies. To this end, I would like to simply close my thesis by listing the least insecure conclusions from our experiments:

1. Perception is a flexible process, allowing for effects of action in a multitude of manners.
2. Ecologically valid biological motion displays are perceived as larger than similar displays that depict ecologically implausible motion cues.
3. On-line effects of action on ambiguous perception are not restricted to a certain level of processing: cognitive effects as well as direct, low-level, connections can both be observed.

6. References

Abraham, H., Reimer, B., Seppelt, B., Fitzgerald, C., Mehler, B, & Coughlin, J. F. (2017). *Consumer interest in automation: Preliminary observations exploring a year's change.* [White paper]. Retrieved January 4, 2018, from MIT Agelab: http://agelab.mit.edu/sites/default/files/MIT%20-%20NEMPA%20White%20Paper%20FINAL.pdf

Ahissar, M. & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences, 8*(10), 457-464.

Althoff, T., Sosic, R., Hicks, J. L., King, A. C., Delp, S. L., & Leskovec, J. (2017). Large-scale physical activity data reveal worldwide activity inequality. *Nature, 547*(7663), 336-339.

Anderson, L. C., Bolling, D. Z., Schelinski, S., Coffman, M. C., Pelphrey, K. A., & Kaiser, M. D. (2013). Sex differences in the development of brain mechanisms for processing biological motion. *Neuroimage, 83,* 751-760.

Arrighi, R., Cartocci, G., & Burr, D. (2011). Reduced perceptual sensitivity for biological motion in paraplegia patients. *Current Biology, 21*(22), 910-911.

Aschersleben, G. & Prinz, W. (1995). Synchronizing actions with events: the role of sensory information. *Perception & Psychophysics, 57*(3), 305-317.

Atkinson, A. P., Tunstall, M. L., & Dittrich, W. H. (2007). Evidence for distinct contributions of form and motion information to the recognition of emotions from body gestures. *Cognition, 104*(1), 59-72.

Bassett, D. R., Wyatt, H. R., Thompson, H., Peters, J. C., & Hill, J. O. (2010). Pedometer-measured physical activity and health behaviors in U. S. adults. *Medicine & Science in Sports & Exercise, 42*(10), 1819-1825.

Beets, I. A. M., Rösler, F., & Fiehler, K. (2010). Nonvisual motor learning improves visual motion perception: Evidence from violating the two-thirds power law. *Journal of Neurophysiology, 104*(3), 1612-1624.

Beets, I. A. M., 't Hart, B. M., Rösler, F., Henriques, D. Y. P., Einhäuser, W., & Fiehler, K. (2010). Online action-to-perception transfer: Only percept-dependent action affects perception. *Vision Research, 50*(24), 2633-2641.

Beintema, J. A. & Lappe, M. (2001). Perception of biological motion without local image motion. *Proceedings of the National Academy of Science, 99*(8), 5661-5663.

Bertenthal, B. I. & Pinto, J. (1994). Global processing of biological motions. *Psychological Science, 5*(4), 221-225.

Barclay, C. D., Cutting, J. E., & Kozlowski, L. T. (1978). Temporal and spatial factors in gait perception that influence gender recognition. *Perception & Psychophysics, 23*(2), 145-152.

Blaker, N. M. & van Vugt, M. (2014). The status-size hypothesis: How cues of physical size and social status influence each other. In J. T. Cheng, J. L. Tracy, & C. Anderson (Eds.), *The Psychology of Social Status* (pp. 119-137). New York: Springer.

Casile, A. & Giese, M. A. (2006). Non-visual motor learning influences the recognition of biological motion. *Current Biology, 16*(1), 69-74.

Chang, D. H. & Troje, N. F. (2009). Acceleration carries the local inversion effect in biological motion perception. *Journal of Vision, 16*(9), 1-17.

Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior, 28*(2), 117-139.

Cutting, J. E. (1981). Coding theory adapted to gait perception. *Journal of Experimental Psychology: Human Perception and Performance, 7*(1), 71-87.

Deen, B. & McCarthy, G. (2010). Reading about the actions of others: Biological motion imagery and action congruency influence brain activity. *Neuropsychologia, 48*(6), 1607-1615.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology, 122*(3), 371-396.

Di Pace, E. & Saracini, C. (2014). Action imitation changes perceptual alternations in binocular rivalry. *PLoS ONE, 9*(5), e98305.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research, 91*(1), 176-180.

Dubois, D., Rucker, D. D., & Galinsky, A. D. (2011). Super size me: Product size as a signal of status. *Journal of Consumer Research, 38*(6), 1047-1061.

Duguid, M. M. & Goncalo, J. A. (2012). Living large: The powerful overestimate their own height. *Psychological Science, 23*(1), 36-40.

Engel, A. K., Maye, A., Kurthen, M., & König, P. (2013). Where's the action? The pragmatic turn in cognitive science. *Trends in Cognitive Sciences, 17*(5), 202-209.

Fagioli, S., Hommel, B., & Schubotz, R. I. (2007). Intentional control of attention: Action planning primes action-related stimulus dimensions. *Psychological Research, 71*(1), 22-29.

Flach, R., Knoblich, G., & Prinz, W. (2004). The two-thirds power law in motion perception. *Visual Cognition, 11*(4), 461-481.

Fox, R. & McDaniel, C. (1982). The perception of biological motion by human infants. *Science, 218*(4571), 486-487.

Geitner, C., Sawyer, B. D., Birrell, S., Jennings, P., Skyrypchuk, L., Mehler, B., & Reimer, B. (2017). *A link between trust in technology and glance allocation in on-road driving.* Paper presented at the Ninth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Manchester Village, Vermont. Retrieved from http://drivingassessment.uiowa.edu/sites/default/files/DA2017/papers/41.pdf

Gibson, J. J. (2015). *The ecological approach to visual perception* (Classic Edition). New York, NY: Taylor & Francis.

Gravano, S., Zago, M., & Lacquaniti, F. (2017). Mental imagery of gravitational motion. *Cortex, 95*,172-191.

Grosjean, M., Shiffrar, M., & Knobilch, G. (2007). Fitts's law holds for action perception. *Psychological Science, 18*(2), 95-99.

Grossman, E. D. & Blake, R. (2001). Brain activity evoked by inverted and imagined biological motion. *Vision Research, 41*(10), 1475-1482.

Hall, C. R., Rodgers, W. M., & Barr, K. A. (1990). The use of imagery by athletes in selected sports. *The Sport Psychologist, 4*(1), 1-10.

Hecht, H., Vogt, S., & Prinz, W., (2001). Motor learning enhances perceptual judgment: a case for action-perception transfer. *Psychological Research, 65*(1), 3-14.

Hemeren, P.E. (2008). *Mind in Action: Action recognition and the perception of biological motion.* Lund University Cognitive Studies 140, Lund University, Sweden.

Hill, H. & Johnston, A. (2001). Categorizing sex and identity from the biological motion of faces. *Current Biology*, *11*(11)*,* 880–885.

Hirai, M., Chang, D. H. F., Saunders, D. R., & Troje, N. F. (2011). Body configuration modulates the usage of local cues to direction in biological-motion perception. *Psychological Science, 22*(12)*,* 1543-1549.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): a framework for perception and action planning. *Behavioral and Brain Sciences, 24*(5), 849-937.

Ikeda, H. & Watanabe, K. (2009). Anger and happiness are linked differently to the explicit detection of biological motion. *Perception, 38*(7), 1002-1011.

Inagaki, T. & Itoh, M. (2013). Human's overtrust in and overreliance on advanced driver assistance systems: A theoretical framework. *International Journal of Vehicular Technology,* vol. 2013, p. 8.

Itto, Y. & Hatta, T. (2004). Spatial structure of quantitative representation of numbers: Evidence from the SNARC effect. *Memory & Cognition, 32*(4), 662-673.

Ivanenko, Y. P., Grasso, R., Macellari, V., & Lacquaniti, F. (2002). Two-thirds power law in human locomotion: Role of ground contact forces. *NeuroReport, 13*(9), 1171-1174.

James, W. (1890). *The principles of psychology.* New York: Holt.

James, W. (1912). *Essays in Radical Empiricism.* New York: Longmans, Green, and Company.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics, 14*(2), 201-211.

Johansson, G. (1976). Spatio-temporal differentiation and integration in visual motion perception. *Psychological Research, 38*(4)*,* 379-393.

Jokisch, D., Daum, I., Suchan, B., & Troje, N. F. (2005). Structural encoding and recognition of biological motion: Evidence from event-related potentials and source analysis. *Behavioral Brain Research, 157*(2), 195-204.

Jokisch, D. & Troje, N. F. (2003). Biological motion as a cue for the perception of size. *Journal of Vision, 3*(4), 252-264.

Jörges, B. & López-Moliner, J. (2017). Gravity as a strong prior: Implications for perception and action. *Frontiers in Human Neuroscience, 11*(203), 1-16.

Keetels, M. & Stekelenburg, J. J. (2014). Motor-induced visual motion: Hand movements driving visual motion perception. *Experimental Brain Research, 232*(9)*,* 2865-2877.

Kidd, D. G., Cicchino, J. B., Reagan, I. J., & Kerfoot, L. B. (2017). Driver trust in five driver assistance technologies following real-world use in four production vehicles. *Traffic Injury Prevention, 18*(1), 44-50.

Knoblich, G. & Flach, R. (2001). Predicting the effects of actions: Interactions of perception and action. *Psychological Science, 12*(6), 467-472.

Kozlowski, L. T. & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, *21*(6)*,* 575–580.

Lacquaniti, F., Carrozzo, M., & Borghese, N. (1993). The role of vision in tuning anticipatory motor responses of the limbs. In A. Berthoz, C. Gielen, V. Henn, K. P. Hoffmann, M. Imbert, F. Lacquaniti & A. Roucoux (Eds), *Multisensory Control of Movement* (pp. 379-393). Oxford: Oxford University Press.

Lacquaniti F., Terzuolo C., & Viviani P. (1983). The law relating the kinematic and figural aspects of drawing movements. *Acta Psychologica, 54*(1), 115-130.

Lange, J. & Lappe, M. (2006). A model of biological motion perception from configural form cues. *Journal of Neuroscience, 26*(11), 2894-2906.

Lindqvist, E. (2012). Height and leadership. *Review of Economics and Statistics, 94*(4), 1191-1196.

Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance, 31*(1), 210-220.

Marsh, A. A., Yu, H. H., Schechter, J. C., & Blair, R. J. R. (2009). Larger than life: Humans' nonverbal status cues alter perceived size. *PLoS ONE, 4*(5), 1-8.

Maruya, K., Yang, E., & Blake, R. (2007). Voluntary action influences visual competition. *Psychological Science, 18*(12), 1090-1098.

Masters, R., Poolton, J., & van der Kamp, J. (2010). Regard and perceptions of size in soccer: better is bigger. *Perception, 39*(9), 1290-1295.

Mather, G. & Murdoch, L. (1994). Gender discrimination in biological motion displays based on dynamic cues. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 258*(1353)*,* 273-279.

Mather, G., Radford, K., & West, S. (1992). Low-level visual processing of biological motion. *Proceedings of the Royal Society London B, Biological Sciences, 249*(1325), 149-155.

Meier, B. P., Robinson, M. D., & Caven, A. J. (2008). Why a big mac is a good mac: Associations between affect and size. *Basic and Applied Social Psychology, 30*(1)*,46-55.*

Meng, M. & Tong, F. (2006). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision, 4*(7), 539-551.

Miller, L. E. & Saygin, A. P. (2013). Individual differences in the perception of biological motion: Links to social cognition and motor imagery. *Cognition, 128*(2), 140-148.

Montepare, J. M., Goldstein, S. B., & Clausen, A. (1987). The identification of emotions from gait information. *Journal of Nonverbal Behavior, 11*(1), 33-42.

Murray, G. R. & Schmitz, J. D. (2011). Caveman politics: Evolutionary leadership preferences and physical stature. *Social Science Quarterly, 92*(5), 1215–1235.

Müsseler, J. (1999). How independent from action is perception? An event-coding account for more equally-ranked crosstalks. In G. Ascherleben, T. Bachman, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 121-147). Amsterdam: Elsevier.

Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. *Nature, 395*(6705), 894-896.

O'Regan, J. K. & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*(5), 939-1031.

Pavlova, M., Krageloh-Mann, I., Sokolov, A., & Birbaumer, N. (2001). Recognition of point-light biological motion displays by young children. *Perception, 30*(8), 925-933.

Pollick, F. E., Lestou, V., Ryu, J., & Cho, S. (2002). Estimating the efficiency of recognizing gender and affect from biological motion. *Vision Research, 42*(20), 2345-2355.

Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition, 82*(2), 51-61.

Polsinelli, M., Milanesi, G., & Ganesan, A. T. (1969). Size adaptation: a new aftereffect. *Science, 166*(3902), 245–247.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology, 9*(2)*,* 129-154.

Reimer, B., Pettinato, A., Fridman, F., Lee, J., Mehler, B., Seppelt, B., Park, J., & Iagnemma, K. (2016). *Behavioral impact of drivers' roles in automated driving.* Paper presented at the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Ann Arbor, MI. Retrieved from http://delivery.acm.org/10.1145/3010000/3005411/p217-reimer.pdf?ip=134.109.16.248&id=3005411&acc=OA&key=2BA2C432AB83DA15%2E2C9D92A4DDBE0279%2E4D4702B0C3E38B35%2E5BD68F0398C58894&CFID=1023817971&CFTOKEN=55585031&__acm__=1515103599_f0f4225348398287c391180ec6d5ccd5

Repp, B. H. & Knoblich, G. (2007). Action can affect auditory perception. *Psychological Science, 18*(1), 6-7.

Saunders, D. R., Suchan, J., & Troje, N. F. (2009). Off on the wrong foot: local features in biological motion. *Perception, 38*(4), 522-532.

Sheikh, A. A. & Korn, E. R. (1994). *Imagery in sports and physical performance.* Amityville, New York: Baywood Publishing Company Inc.

Shi, J., Weng, X., He, S., & Jiang, Y. (2010). Biological motion cues trigger reflexive attentional orienting. *Cognition, 117*(3)*,* 348-354.

Shiffrar, M., Lichtey, L., & Haptulla-Chatterjee, S. (1997). Percepts of biological motion across apertures. *Perception & Psychophysics, 59*(1), 51-59.

Schütz-Bosbach, S. & Prinz, W. (2007). Perceptual resonance: Action-induced modulation of perception. *Trends in Cognitive Sciences, 11*(8), 349-355.

Silvera, D. H., Josephs, R. A., & Giesler, R. B. (2002). Bigger is better: The influence of physical size on aesthetic preference judgments. *Journal of Behavioral Decision Making, 15*(3), 189-202.

Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *PNAS, 105*(2), 809-813.

Sonoda, K. & Wada, T. (2017). Displaying system situation awareness increases driver trust in automated driving. *IEEE Transactions on Intelligent Vehicles, 2*(3), 185-193.

Stephen, D. G., Stepp, N., Dixon, J. A., & Turvey, M. T. (2008). Strong anticipation: Sensitivity to long-range correlations in synchronization behavior. *Physica A: Statistical and Theoretical Physics, 387*, 5271-5278.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276-315.

Sumi, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. *Perception, 13*(3), 283-286.

Thill, S., Hemeren, P. E., & Nilsson, M. (2014). *The apparent intelligence of a system as a factor in situation awareness.* Paper presented at the IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), Orlando, FL. Retrieved from http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6816540

Thornton, I. M. & Vuong, Q. C. (2004). Incidental processing of biological motion. *Current Biology, 14*(12), 1084-1089.

Troje, N. F. (2008). Biological motion perception. In A. I. Basbaum, M. C. Bushnell, D. V. Smith, G. K. Beauchamp, S. J. Firestein, P. Dallos, D. Oertel, R. H. Masland, T. D. Albright, J. H. Kaas, & E. P. Gardner (Eds.), *The Senses: A Comprehensive Reference* (pp. 231-238). Oxford, Elsevier.

Troje, N. F. (2013). What is biological motion?: Definition, stimuli and paradigms. In M. D. Rutherford, & V. A. Kuhlmeier (Eds.), *Social Perception: Detection and Interpretation of Animacy, Agency, and Intention* (pp. 13-36). Cambridge, MA: MIT Press.

Troje, N. F. & Chang, D. H. F. (2013). Shape-independent processes in biological motion perception. In K. L. Johnson & M. Shiffrar (Eds.) *People Watching: Social, Perceptual, and Neurophysiological Studies of Body Perception* (pp. 82-100). New York: Oxford University Press.

Troje, N. F. & Westhoff, C. (2006). The inversion effect in biological motion perception: Evidence for a "life detector"? *Current Biology, 16*(8), 821-824.

Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: Effects of structural and kinematic cues. *Perception & Psychophysics, 67*(4), 667-675.

Tsuchiya, N., Koch, C., Gilroy, L. A., & Blake, R. (2006). Depth of interocular suppression associated with continuous flash suppression, flash suppression, and binocular rivalry. *Journal of Vision, 6*(10), 1068-1078.

Uithol, S., van Rooij, I., Bekkering, H., & Haselager, P. (2011). Understanding motor resonance. *Social Neuroscience, 6*(4), 388-397.

Yap, A. J., Mason, M. F., & Ames, D. R. (2013). The powerful size others down: The link between power and estimates of others' size. *Journal of Experimental Social Psychology, 49*(3), 591-594.

Vallortigara, G., Regolin, L., & Marconato, F. (2005). Visually inexperienced chicks exhibit spontaneous preference for biological motion patterns. *PloS Biology, 3*(7), 1312-1316.

Vanrie, J., Dekeyser, M., & Verfaillie, K. (2004). Bistability and biasing effects in the perception of ambiguous point-light walkers. *Perception, 33*(5), 547–560.

Vanrie, J. & Verfaillie, K. (2004). Perception of biological motion: A stimulus set of human point-light actions. *Behavior Research Methods, Instruments, & Computers, 36*(4), 625-629.

Van Vleet, T. M., Hoang-duc, A. K., DeGutis, J., & Robertson, L. C. (2011). Modulation of non-spatial attention and the global/local processing bias. *Neuropsychologia, 49*(3), 352-359.

Veltkamp, M., Aarts, H., & Custers, R. (2008). Perception in the service of goal pursuit: Motivation to attain goals enhances the perceived size of goal-instrumental objects. *Social Cognition, 26*(6), 720-736.

Veto, P., Einhäuser, W., & Troje, N. F. (2017). Biological motion distorts size perception. *Scientific Reports, 7*(10), 42576.

Veto, P., Schütz, I., & Einhäuser, W. (in press). Continuous flash suppression: Manual action affects eye movements but not the reported percept. *Journal of Vision*

Veto, P., Thill, S., & Hemeren, P. (2013). Incidental and non-incidental processing of biological motion: Orientation, attention and life detection. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.) *Cooperative Minds: Social Interaction and Group Dynamics: Proceedings of the 35th Annual Meeting of the Cognitive Science Society* (pp. 1528-1533). Berlin: Cognitive Science Society.

Veto, P., Uhlig, M., Troje, N. F., & Einhäuser, W. (submitted manuscript). What you see is what you expect: Cognitive assumptions influence the action-to-perception transfer in ambiguous perception.

Wang, L., Yang, X., Shi, J., & Jiang, Y. (2014). The feet have it: Local biological motion cues trigger reflexive attentional orienting in the brain. *NeuroImage, 84*(1), 217-224.

Wang, L., Zhang, K., He, S., & Jiang, Y. (2010). Searching for life motion signals: Visual search asymmetry in local but not global biological-motion processing. *Psychological Science, 21*(8), 1083-1089.

Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology, 52*, 113-117.

Wexler, M., Kosslyn, S. M., & Berthoz, A. (1998). Motor processes in mental rotation.. *Cognition*, *68*(1), 77–94.

Wohlschläger, A. (2000). Visual motion priming by invisible actions. *Vision Research, 40*(8), 925-930.

Wohlschläger, A. & Wohlschläger, A. (1998). Mental and manual rotation. *Journal of Experimental Psychology: Human Perception and Performance, 24*(2), 397-412.

Zacks, J. M. (2004). Using movement and intentions to understand simple events. *Cognitive Science, 28*(6), 979-1008.

Study I.

Biological motion distorts size perception

(2017)

Scientific Reports, 7(10), 42576

Peter Veto[1,2,3], Wolfgang Einhäuser[2], Nikolaus F Troje[3]

[1]Philipps-University Marburg, Department of Physics, Marburg, D-35043, Germany

[2]Chemnitz University of Technology, Institute of Physics, Chemnitz, D-09107, Germany

[3]Queen's University, Department of Psychology, Kingston, ON K7L 3N6, Canada

Abstract

Visual illusions explore the limits of sensory processing and provide an ideal testbed to study perception. Size illusions – stimuli whose size is consistently misperceived – do not only result from sensory cues, but can also be induced by cognitive factors, such as social status. Here we investigate, whether the ecological relevance of biological motion can also distort perceived size. We asked observers to judge the size of point-light walkers (PLWs), configurations of dots whose movements induce the perception of human movement, and visually matched control stimuli (inverted PLWs). We find that upright PLWs are consistently judged as larger than inverted PLWs, while static point-light figures did not elicit the same effect. We also show the phenomenon using an indirect paradigm: observers judged the relative size of a disc that followed an inverted PLW larger than a disc following an upright PLW. We interpret this as a contrast effect: The upright PLW is perceived larger and thus the subsequent disc is judged smaller. Together, these results demonstrate that ecologically relevant biological-motion stimuli are perceived larger than visually matched control stimuli. Our findings present a novel case of illusory size perception, where ecological importance leads to a distorted perception of size.

Introduction

Systematic distortions in the perception of size can be observed in a wide variety of visual scenarios. Two mechanisms underlie most of the classic examples. One of them is size constancy, where an object that appears farther from the viewer seems to be larger as opposed to a nearer object, even though they create an equally large retinal image in the viewer. The other mechanism is size contrast, where the apparent size of an object changes inversely with the size of other, related, objects. This can take place simultaneously (e.g., a circle among circles in the Ebbinghaus/Titchener illusion), or with a temporal delay (size adaptation aftereffect[1]). Illusions exploiting these mechanisms affect not only the "conscious" percept as reported by the viewer, but also the size of afterimages[2] or objective measures, such as reaction times[3-4].

While the aforementioned size illusions are perceptual in nature, a different class of size illusions pertains to social constructs that can also lead to a change in perceived size of a person or an inanimate object. A general association between positive subjective value and larger size exists[5], and this reciprocal connection has been observed in different areas of life. The most palpable example for such a relation is between social leadership and physical size[6], where it is conspicuous that mechanisms described by evolutionary psychology still play a role in today's society[7]. Tall men are more likely to take managerial positions than short men[8], while people with more social power perceive other humans[9] and objects[10] as smaller, as well as they are perceived as taller by others[11] and by themselves[10]. A size-status connection also prevails in the case of consumer products[12]. Aside from power, motivation and action goals[13] and aesthetic preference judgments[14] are likewise related to the size of non-animate objects. Altogether, these findings suggest that there is a general, positive, association between the importance or value of an object to the viewer and its perceived size.

Animate motion patterns are rapidly perceivable[15,16,17], visually salient[18] and carry numerous types of information that are readily retrievable by human observers[19,20,21]. The perception of biological motion is arguably of high ecological importance, making preferential processing by the visual system for such stimuli likely, even though direct evidence is scarce. The most commonly used tool to explore this question is point-light figures. They eliminate all visual information obtainable from the surface of the body by only showing the movements of a few important articulations depicted as dots. With the help of point-light stimuli, biological motion has been shown to yield to several perceptual benefits as compared to similar non-biological motion. For example, coherent and upright point-light walkers (PLWs) are processed incidentally in a flanker paradigm, as opposed to static, scrambled[22] or inverted[23] walkers. Upright, scrambled biological motion stimuli lead to faster hits in a search task than similar, but inverted figures[24], which means that local cues of biological motion act on a preattentive level of visual processing. Upright human or terrestrial animal PLWs induce reflexive attentional orienting in a central cueing paradigm, while inverted or static figures do not[25], showing that incidental effects are not specific to stimuli presenting configural information that is typical of humans. Indeed, local motion cues, in particular those of the feet, play a crucial role in a “life detector” system: a general filter in human vision, tuned to help us detect terrestrial animals[26-28].

Biological-motion stimuli, therefore, seem to be of special importance in visual processing. Also, important objects tend to look larger to the viewer. We thus hypothesize that stimuli carrying ecologically valid biological motion cues appear larger to observers, as compared to similar motion stimuli lacking ecological validity. We tested this hypothesis using human PLWs in three experiments, where we compared coherent, upright, PLW figures to inverted ones. In the inverted displays, both local and global biological motion cues lack ecological validity, while all other aspects of the stimulus remain equal to those in the upright figures. Hence, seeing a difference in perceived size between the two conditions can only be due to the effect of the ecological importance of biological motion.

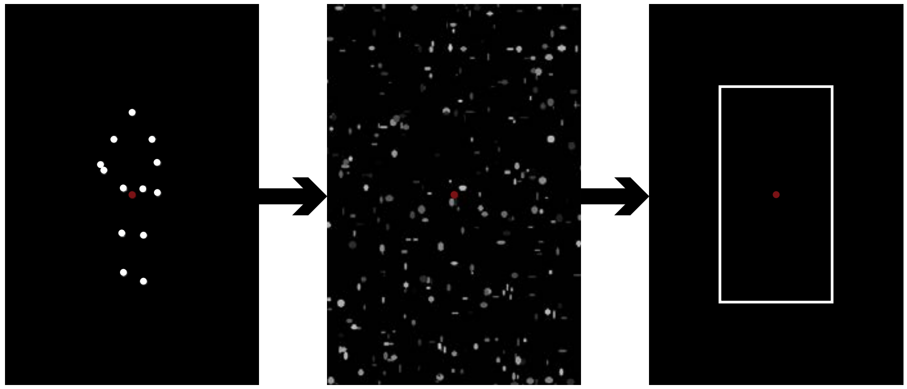

Figure 1. Paradigm – Experiment I. Sequence of a single trial (here with upright PLW).

Experiment I

Perceived sizes of upright and inverted PLWs (see Figure 1) were compared directly in an adjustment task.

Methods

Participants. Sixteen students from the Queen's University participant pool (one male, fifteen females, mean age = 20.1, *SD* = 1.8) participated in the study. Experimental protocols of all experiments conformed to the World Medical Association Declaration of Helsinki and were approved by the board "Ethikkommission FB04, Philipps-University Marburg" and by the Human Ethics committee at Queen's. All participants had normal or corrected-to-normal vision, provided written informed consent and received monetary compensation. One participant dropped out after reporting problems with larger stimulus sizes.

Stimuli. Upright and inverted PLWs were depicted from a frontal view, based on the action "Walk" from a stimulus set of human actions created by Vanrie and Verfaillie[29], based on the actions of a male actor. The figure consisted of 13 dots, showing the positions of the head and the main articulations of the limbs (Figure 1). Walker size was varied in 10 steps between 2.44˚ x 0.88˚ and 7.86˚ x 2.70˚ (mean:

5.24° x 1.77°) at a viewing distance of 75 cm. Each PLW presentation started at a random frame of the stride, resulting in slightly varying sizes for each trial. All displays were gray on black background, with a red fixation point continuously shown in the center of the screen. Stimuli were presented on a 17” CRT screen with Matlab and the Psychophysics Toolbox[30,31].

Procedure. For each trial, participants were asked to maintain fixation on the fixation point and viewed a centrally displayed PLW for 250 ms, followed by a dynamic random dot mask, lasting for 200 ms. After the mask, participants had to move the mouse in order to adjust a rectangle to frame the area occupied by the previously seen walker as tightly as possible (Figure 1). The mouse position was connected to a corner of the rectangle, starting randomly either from the fixation point, or from well outside of the stimulus’ area. The rectangle stayed centrally symmetrical at all times. That way, the width and height of the walker were set independently, albeit in a single response. Participants confirmed their responses by a mouse click, after which the next trial started following a random intertrial interval between 500 and 800 ms. Each participant completed 400 trials.

Analysis. For each trial, the percentage of overestimation (area of the response rectangle divided by the area of the smallest frame containing all dots at any time) was calculated. Outlier responses (cutoff = 2.5 SD) were removed for each block (2.2% of all trials). A one sample t-test was carried out to determine whether the difference between responses to upright and inverted walkers (Distortion Effect = $\text{Overestimation}_{\text{Upright Walker}} - \text{Overestimation}_{\text{Inverted Walker}}$) was significantly different from zero.

Results

The size distortion effect was significantly different from zero (expressed in percentage of walker area: mean = 9.07, $SD = 5.73$; $t(14) = 6.12$, $p < 0.001$). This confirms our hypothesis that upright walkers are perceived to be larger than inverted walkers.

Experiment II

To control whether the observed size-distortion effect is specific to biological *motion* (rather than an upright/inverted difference per se), we conducted a second experiment similar to Experiment I, with the additional condition of static point-light figures. If the effect is caused by configural information alone instead of biological motion, static figures should elicit the same pattern of results as dynamic PLWs.

Methods

Participants. Twenty-four students from the Chemnitz University of Technology (five males, nineteen females, mean age = 21.9, *SD* = 3.2) participated in the study.

Stimuli. Stimuli were presented on a 23.6" screen (VPixx Technologies Inc., Saint-Bruno, QC Canada), with all other details of the stimulus kept equal to those in Experiment I. In each trial of the additional static condition, a randomly selected frame of the PLW was presented for the same duration of time (250 ms) as the moving PLW in the dynamic condition.

Procedure. Each participant completed a total of 640 trials split over four blocks. Two blocks contained dynamic PLWs while the other two contained static point-light figures. Static and dynamic trials were otherwise identical. The order of the four blocks was counterbalanced across observers.

Results

Dynamic blocks showed a replication of results from Experiment I, with a size distortion effect significantly different from zero (in percentage of walker area: mean = 7.34, $SD = 15.19$; $t(23) = 2.37$, $p = 0.03$). Static blocks on the other hand did not show a significant size-distortion effect (mean = 4.32, $SD = 17.60$; $t(23) = 1.20$, $p = 0.24$).

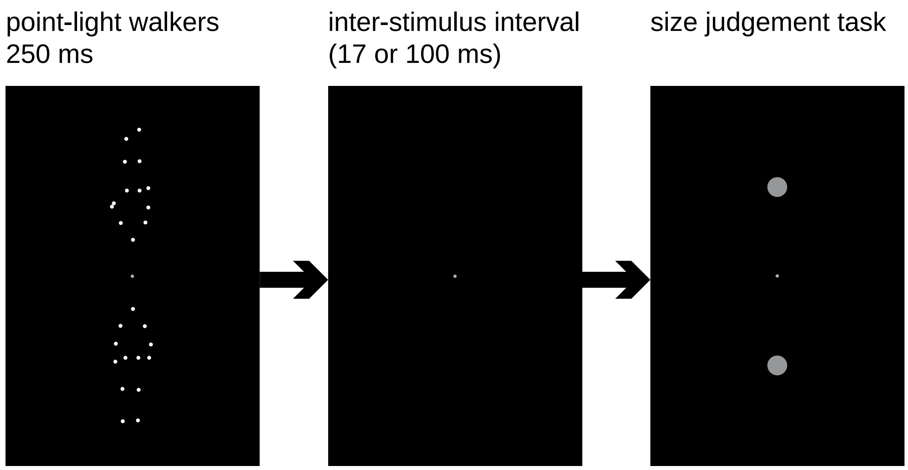

Figure 2. Paradigm – Experiment III. Sequence of a single trial (here with upright PLW in the lower position and inverted PLW in the upper position).

Experiment III

Perceived sizes of upright and inverted PLWs were compared indirectly, with a size judgment task on targets presented subsequently to PLWs. As upright PLWs are perceived as larger, we expect that contrast effects will lead to a subsequent target to appear as smaller. Since participants react to simple disc targets and they are instructed to ignore the preceding figures, this experiment further ensures that our previous findings are caused by a perceptual distortion of size and not by any unexplored bias related to PLWs.

Methods

Participants. Sixteen students (five males, eleven females, mean age = 22.1, *SD* = 1.9) participated in the study. Eight (1-8) were measured at Philipps-University Marburg and eight (9-16) at Queen's University, and recruited through the respective participant pools. All participants had normal or corrected-to-normal vision, provided written informed consent and received monetary compensation.

Stimuli. Generation and presentation of stimuli were as described for Experiment I. PLWs (both walkers, in all conditions: 5.4° x 1.9°) and target discs (diameters depending on condition: 0.76° &

0.76˚; 0.72˚ & 0.80˚; 0.68˚ & 0.84˚) were presented centered 3.46˚ above and below fixation. All displays were gray on black background, and a fixation point was continuously shown in the center of the screen.

Procedure. For each trial, participants were asked to maintain fixation on the fixation point while viewing two PLWs (one upright and the other inverted) for 250 ms. Participants were instructed to ignore these displays. Following a blank inter-stimulus interval (ISI) of 17 or 100 ms, two target discs appeared for 100 ms at the locations of the previously seen walkers (Figure 2). Targets were either identical or differed in size (10.5% or 21% larger or smaller than the average size of 0.76˚). Participants gave a non-speeded forced choice response by pressing one of two buttons, indicating which of the targets was larger than the other. After response, the next trial started following a random intertrial interval between 500 and 800 ms.

Participants 1-8 also completed trials for a temporal judgment task in separate blocks, which are not reported here. For participants 1-4, no trials with identical targets were presented. For participants 5-8, eye tracking was used to validate that observers maintained fixation throughout stimulus presentation. Participants 1-4 each performed 400 trials, participants 5-8 each performed 480 trials and participants 9-16 each performed 1000 trials in total.

Analysis. For each participant, the point of subjective equality (PSE) between targets preceded by upright and inverted walkers was calculated. To do so, a psychometric function was fitted to the data of each individual (fraction of responses "larger" at upright PLW location vs. size difference of discs), and the PSE determined analytically from its two fit parameters (cf. Figure 3). A one-sample t-test was then used to determine whether PSEs were significantly different from zero.

Results

PSEs were shifted towards larger targets at the upright PLW's location (mean = 2.60, *SD* = 2.96, in percentage of target size). This shift was different from zero ($t(15) = 3.51$, $p = 0.003$). There was no difference between trials with long and short ISIs ($mean_{Short\ ISI}$ = 2.63, $SD_{Short\ ISI}$ = 3.26; $mean_{Long\ ISI}$ = 2.68, $SD_{Long\ ISI}$ = 3.44; $t(15) = 0.06$, $p = 0.95$). This is in line with our hypothesis and shows that targets preceded by an upright walker are perceived as smaller than targets preceded by an inverted walker (Figure 3).

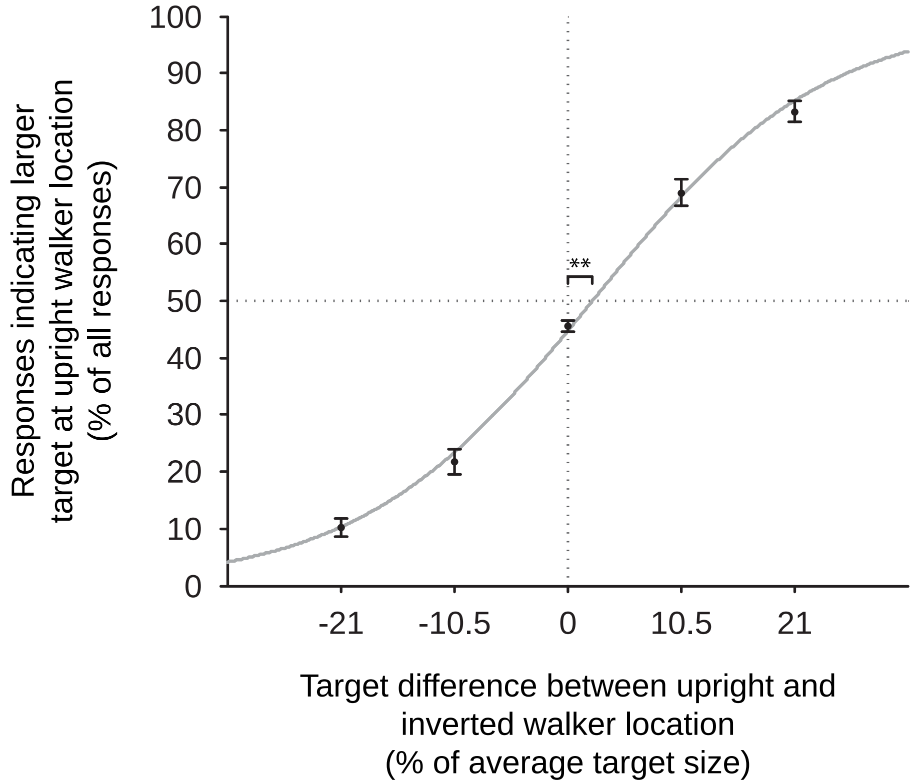

Figure 3. Results – Experiment III. Percent of responses indicating that the target preceded by an upright walker was larger plotted against the difference between target (disc) sizes. Means per condition with fitted psychometric function. Error bars show s.e.m. Asterisks indicate significant difference of PSE at $p < 0.01$. Data for both ISIs (17 ms, 100 ms) were aggregated for analysis.

The functional form of the psychometric function is given by f(x;a,l)=a/(a+exp(-l*x)), and thus the PSE by x=-ln[a]/l with fit parameters a and l. Note that the psychometric function for illustration is a fit to the average data, while for statistical analysis each individual was fitted with a separate psychometric function and analysis was based on the distribution of the individual PSEs.

Discussion

The findings presented here show that stimuli with ecologically valid biological motion cues appear larger than similar motion stimuli without ecological validity. Experiment I demonstrates this phenomenon. Experiment II replicates the findings and shows that static point-light displays do not lead to a similar distortion in perceived size. Experiment III shows that the effect can also be measured indirectly, as it extends through a contrast mechanism to subsequently presented, neutral, stimuli.

Prior studies have demonstrated that discrimination of biological motion stimuli takes place at an early stage of visual processing[16,17] and induces reflexive attentional orienting[25]. This suggests that biological motion stimuli bear high importance, which is further supported by experiments demonstrating that humans[32] and other animals[33] have an innate sensitivity to visual invariants characteristic to biological motion. Our findings lead to similar conclusions, as already a brief presentation (250 ms) of biological motion results in a positive distortion of perceived size, which is linked to subjectively important stimuli[5-14].

Although a contrast effect seems the most likely mechanism transferring the distortion in perceived size from PLWs to the disc targets used in Experiment III, alternative causes are also possible. For example, spatial attention might be deployed asymmetrically between upright and inverted walkers, causing an inhibition of return[34] on responses to subsequent target discs. This, however, would not explain the results found in Experiment I & II, where only one, central target is presented at a time.

While PLWs are useful in eliminating surface information from the body, they thus also take biological motion cues out of their natural context. We cannot exclude that from the dots of a point-light figure the perceptual system might “fill in” the rest of the body. If that happens more likely for upright than for inverted figures, a larger percept would be formed for the former. However, Experiment II offers some hint that this may not be the case in our experiments, as the human figure is also clearly recognizable from the frontal view of a static point-light display (cf. Figure 1).

We cannot exclude that sex differences might also play some role in the results, considering that our participant population was dominantly females and it is conceivable that women are more responsive to biological motion and its social implications[35]. Studies on the link between social power and size[6-11] suggest that the sex as well as the displayed power of the stimulus figure can likewise affect the outcome. Studying sex differences of the reported effects might therefore be an interesting extension in further research.

As it has been shown with other stimuli already, importance to the viewer makes objects look larger. Our data show that biological motion is no exception. It clearly demonstrates a so far unknown example of distorted size perception. Unlike previous examples, this phenomenon is neither a low-level effect[1-4] nor based on social constructs[5,7,9,10]. Instead, our data suggest that the ecological relevance of a biological motion stimulus makes it incidentally appear larger.

References

1. Polsinelli, M., Milanesi, G. & Ganesan, A. T. Size adaptation: A new aftereffect. *Science*, **166**, 245-247 (1969).
2. Sperandio, I., Lak, A. & Goodale, M. A. Afterimage size is modulated by size-contrast illusions. *J. Vis.* **12**, 1-10 (2012).
3. Sperandio, I., Savazzi, S. & Marzi, C. A. Is simple reaction time affected by visual illusions? *Exp. Brain Res.* **201**, 345-350 (2010).
4. Savazzi, S., Emanuele, B., Scalf, P. & Beck, D. Reaction times and perceptual adjustments are sensitive to the illusory distortion of space. *Exp. Brain Res.* **218**, 119-128 (2012).
5. Meier, B. P., Robinson, M. D. & Caven, A. J. Why a big mac is a good mac: Associations between affect and size. *Basic Appl. Soc. Psych.* **30**, 46–55 (2008).
6. Murray, G. R. & Schmitz, J. D. Caveman politics: Evolutionary leadership preferences and physical stature. *Soc. Sci. Q.* **92**, 1215–1235 (2011).

7. Blaker, N. M. & van Vugt, M. The status-size hypothesis: How cues of physical size and social status influence each other in *The Psychology of Social Status* (ed. Cheng, J. T., Tracy, J. L. & Anderson, C.), 119–137 (Springer, 2014).
8. Lindqvist, E. Height and leadership. *REStat* **94**, 1191-1196 (2012).
9. Yap, A. J., Mason, M. F. & Ames, D. R. The powerful size others down: The link between power and estimates of others' size. *J. Exp. Soc. Psychol.* **49**, 591-594 (2013).
10. Duguid, M. M. & Goncalo, J. A. Living large: The powerful overestimate their own height. *Psychol. Sci.* **23**, 36-40 (2012).
11. Marsh, A. A., Yu, H. H., Schechter, J. C. & Blair, R. J. R. Larger than life: Humans' nonverbal status cues alter perceived size. *PLoS ONE* **4**, 1-8 (2009).
12. Dubois, D., Rucker, D. D. & Galinsky, A. D. Super size me: Product size as a signal of status. *J. Cons. Res.* **38**, 1047-1061 (2011).
13. Veltkamp, M., Aarts, H. & Custers, R. Perception in the service of goal pursuit: Motivation to attain goals enhances the perceived size of goal-instrumental objects. *Soc.Cogn.* **26**, 720-736 (2008).
14. Silvera, D. H., Josephs, R. A. & Giesler, R. B. Bigger is better: The influence of physical size on aesthetic preference judgments. *J. Behav. Decis. Mak.* **15***,* 189-202 (2002).
15. Johansson, G. Spatio-temporal differentiation and integration in visual motion perception. *Psychol Res*, **38**, 379-393 (1976).
16. Jokisch, D., Daum, I., Suchan, B. & Troje, N. F. Structural encoding and recognition of biological motion: Evidence from event-related potentials and source analysis. *Behav. Brain Res.*, **157**, 195–204 (2005).
17. Wang, L., Yang, X., Shi, J. & Jiang, Y. The feet have it: Local biological motion cues trigger reflexive attentional orienting in the brain. *NeuroImage*, **84**, 217-224 (2014).
18. Tyler, S. C. & Grossman, E. D. Feature-based attention promotes biological motion recognition. *JoV*, **11**, 1-16 (2011).
19. Barclay, C. D., Cutting, J. E. & Kozlowski, L. T. Temporal and spatial factors in gait perception that influence gender recognition. *Percept. Psychophys.*, **23**, 145-152 (1978).
20. Montepare, J. M., Goldstein, S. B. & Clausen, A. The identification of emotions from gait information. *J. Nonverbal Behav.*, **11**, 33-42 (1987).
21. Troje, N. F., Westhoff, C. & Lavrov, M. Person identification from biological motion: Effects of structural and kinematic cues. *Percept. Psychophys.*, **67**, 667-675 (2005).

22. Thornton, I. M. & Vuong, Q. C. Incidental processing of biological motion. *Curr. Biol.* **14***,* 1084-1089 (2004).

23. Veto, P., Thill, S. & Hemeren, P. Incidental and non-incidental processing of biological motion: Orientation, attention and life detection in *Cooperative Minds: Social Interaction and Group Dynamics: Proceedings of the 35th Annual Meeting of the Cognitive Science Society* (ed. Knauff, M., Pauen, M., Sebanz, N. & Wachsmuth, I.) 1528-1533 (Cognitive Science Society, 2013).

24. Wang, L., Zhang, K., He, S. & Jiang, Y. Searching for life motion signals: Visual search asymmetry in local but not global biological-motion processing. *Psychol. Sci.* **21**, 1083-1089 (2010).

25. Shi, J., Weng, X., He, S. & Jiang, Y. Biological motion cues trigger reflexive attentional orienting. *Cogn.* **117***,* 348-354 (2010).

26. Hirai, M., Chang, D. H. F., Saunders, D. R. & Troje, N. F. Body configuration modulates the usage of local cues to direction in biological-motion perception. *Psychol. Sci.* **22***,* 1543-1549 (2011).

27. Troje, N. F. & Chang, D. H. F. Shape-independent processes in biological motion perception in *People Watching: Social, Perceptual, and Neurophysiological Studies of Body Perception* (ed. Johnson, K. L. & Shiffrar, M.) 82-100 (Oxford University Press, 2013).

28. Troje, N. F. & Westhoff, C. The inversion effect in biological motion perception: Evidence for a "life detector"? *Curr. Biol.* **16***,* 821-824 (2006).

29. Vanrie, J. & Verfaillie, K. Perception of biological motion: A stimulus set of human point-light actions. *Behav. Res. Methods Instrum. Comput.* **36**, 625-629 (2004).

30. Brainard, D. H. The Psychophysics Toolbox. *Spatial Vis.* **10**, 433- 436 (1997).

31. Pelli, D. G. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vis.* **10**, 437-442 (1997).

32. Simion, F., Regolin, L. & Bulf, H. A predisposition for biological motion in the newborn baby. *PNAS*, **105**, 809-813 (2008).

33. Vallortigara, G., Regolin, L. & Marconato, F. Visually inexperienced chicks exhibit spontaneous preference for biological motion patterns. *PloS Biol.*, **3**, 1312-1316 (2005).

34. Posner, M. I. & Cohen, Y. Components of visual orienting in *Attention and Performance Vol. X* (ed. Bouma, H. & Bouwhuis, D.) 531–556 (Erlbaum, 1984).

35. Anderson, L. C., Bolling, D. Z., Schelinski, S., Coffman, M. C., Pelphrey, K. A. & Kaiser, M. D. Sex differences in the development of brain mechanisms for processing biological motion. *Neuroimage*, **83**, 751-760 (2013).

Acknowledgements

This project was supported by the German Research Foundation, International Research Training Group, IRTG 1901, "The Brain in Action" and an NSERC Discovery grant to NFT. The authors are also grateful to Zak Zaman, Paul Hemeren, Doug Munoz, Brian White and Martin Lellep for their ideas and helpful comments.

Study II.

Continuous flash suppression: Manual action affects eye movements but not the reported percept

(in press)

Journal of Vision

Peter Veto[1,2], Immo Schütz[2] & Wolfgang Einhäuser[2]

[1)]Department of Physics, Philipps-Universität, Marburg, Germany

[2)]Institute of Physics, Chemnitz University of Technology, Chemnitz, Germany

Abstract

Diverse paradigms, including ambiguous stimuli and mental imagery, have suggested a shared representation between motor and perceptual domains. We examined the effects of manual action on ambiguous perception in a continuous flash suppression (CFS) experiment. Specifically, we asked participants to try and perceive a suppressed grating, while rotating a manipulandum. In one condition, the grating's motion was fully controlled by the manipulandum movement, in another condition the coupling was weak, and in a third condition no movement was executed. We found no effect of the movement condition on the subjectively reported visibility of the grating, which is in contrast to previous studies that allowed for more top-down influence. However, we *did* observe an effect on eye movements: the gain of the optokinetic nystagmus induced by the grating was modulated by its coupling to the manual movement. Our results (i) indicate that action-to-perception transfer can occur on different levels of perceptual organization, (ii) support the notion that CFS operates qualitatively differently from other ambiguous stimuli, including binocular rivalry, and (iii) highlight the importance of objective measures beyond subjective report when studying how action affects perception and awareness.

Keywords: event-coding, common coding theory, continuous flash suppression, action-perception coupling, action-to-perception transfer, eye-movement, optokinetic nystagmus, ambiguous perception, bistable perception, rivalry, binocular rivalry

Introduction

Perception obviously affects action, but there is also mounting evidence for the reverse direction, a direct effect of action on perception and perceptual representations. In an effort of determining the underlying mechanisms, the common coding theory (Prinz, 1997) suggests that late stages of perception and early stages of action use shared representations, which then allow action planning to be promptly influenced by a connected percept. Furthermore, while an external influence of action on perception is trivial (by moving our eyes, changing location or manipulating the world around us; see e.g., Wexler & van Boxtel, 2005), the common coding theory, as well as the theory of event-coding (e.g. Müsseler, 1999), also implies an internal influence of action on perception. If the observation of an action creates motor resonance, the production of an action should lead to a similar, but perceptual, resonance (Schütz-Bosbach & Prinz, 2007). There is some experimental evidence to support this notion. Wohlschläger and Wohlschläger (1998) found that concurrent manual rotation led to faster performance in a mental rotation task when the directions of action and mental rotation were congruent as opposed to when they were incongruent. This, however, was only observed when the two rotations occurred about the same axis. Similarly, Wexler, Kosslyn and Berthoz (1998) showed that speed as well as accuracy in an imagery mental rotation task can be enhanced by unseen motor rotation, in a direction that is congruent with the action. In line with Wohlschläger and Wohlschläger (1998), they also found that the facilitating effect of congruent action is dependent on the relative angle and velocity of the movements.

Further evidence for an internal effect of action on perception is provided by paradigms in which action leads to a bias in the percept of an external stimulus, instead of the outcome of imagery (for an overview of different methods, see Zwickel & Prinz, 2012). Multistable stimuli are ideal tools for making internal biases measurable, while the external stimulation stays unchanged. This can be achieved through an inherent ambiguity, where the stimulus itself has two or more stable interpretations (ambiguous stimuli such as – for example – the Necker cube (Necker, 1832) or Boring's old/young figure (Boring, 1930)), or through the presentation of different stimuli in one eye as opposed to the other, leading to a conflict that is resolved by either one or the other eye gaining dominance at a time (binocular rivalry, BR, Wheatstone, 1838).

Using an ambiguous rotating display, Wohlschläger (2000) primed the perceived direction of motion by both hand movements and planned actions, and found that either one can be sufficient for biasing the percept, if they share a common, cognitively specified dimension with the stimulus. In the case of planned actions, there is no kinesthetic information to bias the visual modality. He argues that this makes a direct, low level, motor-to-visual information flow an implausible explanation for why the priming – or "action capture" – took place. Instead, he suggests that the phenomenon is more similar to attentional capture, where top-down control plays a crucial role, and where effects are object- and action-centered.

For an ambiguous structure-from-motion cylinder, Beets and colleagues (2010) demonstrated that rotating a manipulandum congruently with the current percept stabilizes its perceptual state, while an opposing rotation yields destabilization. In line with Wohlschläger's reasoning for the top-down nature of action capture, these effects were only observed when the manipulandum was used to report the percept, while no effect of action was found when concurrent rotation of the manipulandum was unrelated to the task.

Instead of ambiguous displays, Keetels and Stekelenburg (2014) used an unambiguous stimulus with high perceptual uncertainty. With a flashing bar stimulus that was displaced only slightly (or remained stationary) at each trial, they found that concurrent, directional, button press actions shifted the point of subjective equality of perceptual displacement judgments in the direction of action.

While both Wohlschläger (2000) and Beets *et al.* (2010) used ambiguous displays, similar effects were also found in BR (Maruya, Yang, & Blake, 2007), where perception of the movement of one eye's stimulus was positively biased by manual control. Interestingly, the effect of action did not only occur when the coupled stimulus was dominant, but also when it was suppressed from awareness (i.e., when the stimulus of the other eye had exclusive dominance). This argues against an explanation that is solely based on top-down control.

Imitated hand action can also bias perception in BR. Di Pace and Saracini (2014) used a dynamic hand action presented in one eye and a checkerboard pattern in the other, and found that dominance durations for perceiving the hand action were longer, when the same action was imitated by the observer.

In a stream-bounce display (Metzger, 1934), two identical objects move towards each other until reaching a common position, after which they move away from each other following a continuous path. The perceptual interpretation of this animation is ambiguous, as the objects may seem to either pass through or bounce off of one another. When the motion of the disks is controlled by the hand action of the observer, the visual interpretation that is congruent with the performed action is more likely to take place, as shown by Mitsumatsu (2009). Similarly to Beets and colleagues (2010), he also found that the mere presence of action is not sufficient for the effect to occur.

Perceptual resonance may happen simultaneously (on-line effects), or through motor learning on different time scales (off-line effects; Schütz-Bosbach & Prinz, 2007). To our knowledge, only a few studies on the on-line effects of action on perception so far have used action as a way of controlling concurrent stimulus dynamics, and they only used action as either coupled or not coupled to the stimulus. To test whether the action-perception transfer is influenced by the degree of coupling between task-relevant hand movement and stimulus velocity dynamics, we used a continuous flash suppression (CFS, Tsuchiya & Koch, 2005) paradigm, where a faint, moving grating stimulus, presented to one eye, was set against a colorful Mondrian suppressor that was displayed to the other. This paradigm constitutes a highly biased variant of inter-ocular conflict, where the strength of one stimulus (the suppressor) is maximal and the other stimulus can only be perceived when endogenous attention is deployed. We varied the degree of coupling between participants' rotational hand action and grating stimulus dynamics (fully coupled action, partially coupled action, no action). Observers were asked to report on their subjective percept by pressing and holding a button whenever they perceived the grating. In addition to this subjective measure that is prone to subjective criterion and response bias, and might only reveal percepts that are clearly suprathreshold, we used an additional measure, which might reveal effects on the suppressed stimulus: throughout the experiment we measured eye position. When the grating becomes dominant, we expect the drifting grating to induce an optokinetic nystagmus (OKN). Unlike the button press, which is an all-or-none report, the gain of the OKN slow phase should be related to the degree of dominance (cf. Naber, Frässle & Einhäuser, 2011, for the case of BR). Hence, we expect that measuring the gain may reveal subtle changes in the visual representation of the grating, even if it is still subjectively suppressed from perception. Consequently, we use the OKN gain as a measure of whether concurrent action influences the perceptual representation of the grating, both above and below perceptual threshold.

This paradigm is particularly applicable for studying the effects of action on perception, as only one stimulus needs to be coupled with the action, while the other remains constant at all times. Since in our case, constant action is required, linear hand movements would not be suitable. For eliciting OKN, however, the linear translation of the stimulus is necessary. The coupling between rotational input and linear output is always arbitrary: as with a rack and pinion mechanism, the direction of coupling depends on the relative spatial configuration of the machine. This renders a traditional BR paradigm with two gratings of opposing directions not applicable for our purposes, as coupling the manipulandum rotation to one stimulus would also result in an equal coupling to the other.

As a form of interocular conflict, CFS has close resemblance to BR, although it shows dissimilarities beyond the greater exerted strength of suppression (Tsuchiya & Koch, 2005; Tsuchiya, Koch, Gilroy & Blake, 2006). As detailed above, several cases of dynamic multistable stimuli have been reported where action influences perception. Based on these results, we hypothesized an increase in perceived strength of the grating when action is coupled to the stimulus' dynamics, as compared to when the same visual stimuli are presented without action. The type of measure that would grasp the increased perceptual strength of the grating stimulus is not yet entirely clear, as CFS has so far not been used extensively in similar settings. However, Levelt's propositions (Levelt, 1965) appear to frequently apply also outside of binocular rivalry paradigms (Brascamp, Klink & Levelt, 2015). Based on these observations, we can hypothesize that larger perceived strength would be expressed in an increase of the grating's predominance (dominance as proportion of time throughout the experiment) and in a decrease of the suppressor's dominance duration (mean of all the individual dominance periods). If viewers perceive the two stimuli as relatively balanced in strength, an increase in the dominance duration for the grating stimulus can also be predicted; furthermore, such a change would increase the alternation rate between the two stimuli. Besides and in parallel with these changes in subjective reports of participants, we also expect to see an increased OKN response to the grating stimulus when it is coupled to manual action as opposed to when no action is performed.

Results from the partially coupled action condition of the experiment will show whether the action-to-perception transfer depends on the type of coupling. When action is coupled directly to the grating's movement, the change in stimulus position is predictable from action. When action is only partially coupled, the action remains task-relevant and still has an effect on stimulus dynamics; however, exact stimulus parameters cannot be estimated. If such perceptual estimates are not necessary for the action-to-perception transfer to take place, we would see similar results with partially coupled action as with full coupling. If, on the other hand, matching dynamics between action and perception do play a role in the expected facilitation, results will be either identical to when no action is required, or somewhere between the no action and fully coupled action conditions.

Methods

Participants

Twenty-four students (15 female, 9 male, 4 left handed, mean age = 24.4 *SD* = 5.8) took part in the experiment. All of them were naïve to the purpose of the experiment, had normal or corrected-to-normal vision and normal stereo vision, and gave written informed consent to their partaking. All procedures conformed to the Declaration of Helsinki and were approved by the applicable board ("Ethikkommission FB04, Philipps-University Marburg").

Setup and stimuli

Manual action was tracked by a custom-built manipulandum device (Fig. 1). The manipulandum consisted of a horizontal disk of 9 cm radius, which had an 11.5 cm long handle mounted perpendicular to the disk at 5 cm distance from the center. To track the angular position of the handle, the axis of the disk operated a Kübler Sendix 5020 incremental rotary encoder. Stimuli were generated using Matlab (Mathworks, Natick, MA) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Presentation took place at a viewing distance of 30 cm on two 21-inch Syncmaster CRT screens (Samsung, Seoul, South Korea), at 1280 × 1024 pixels spatial and 85 Hz temporal resolution. A mirror stereoscope with cold (infrared-transparent) mirrors (Naber *et al*., 2011) allowed for simultaneous dichoptic stimulus presentation and noninvasive infrared eye-tracking of one eye at 1000 Hz (EyeLink 2000, SR Research, Ottawa, ON, Canada with the Eyelink toolbox, Cornelissen, Peters & Palmer, 2002). Since the eye-movement characteristics that we needed to analyze for the purpose of this study are equitably carried by both eyes (Naber *et al*., 2011), binocular tracking was not necessary.

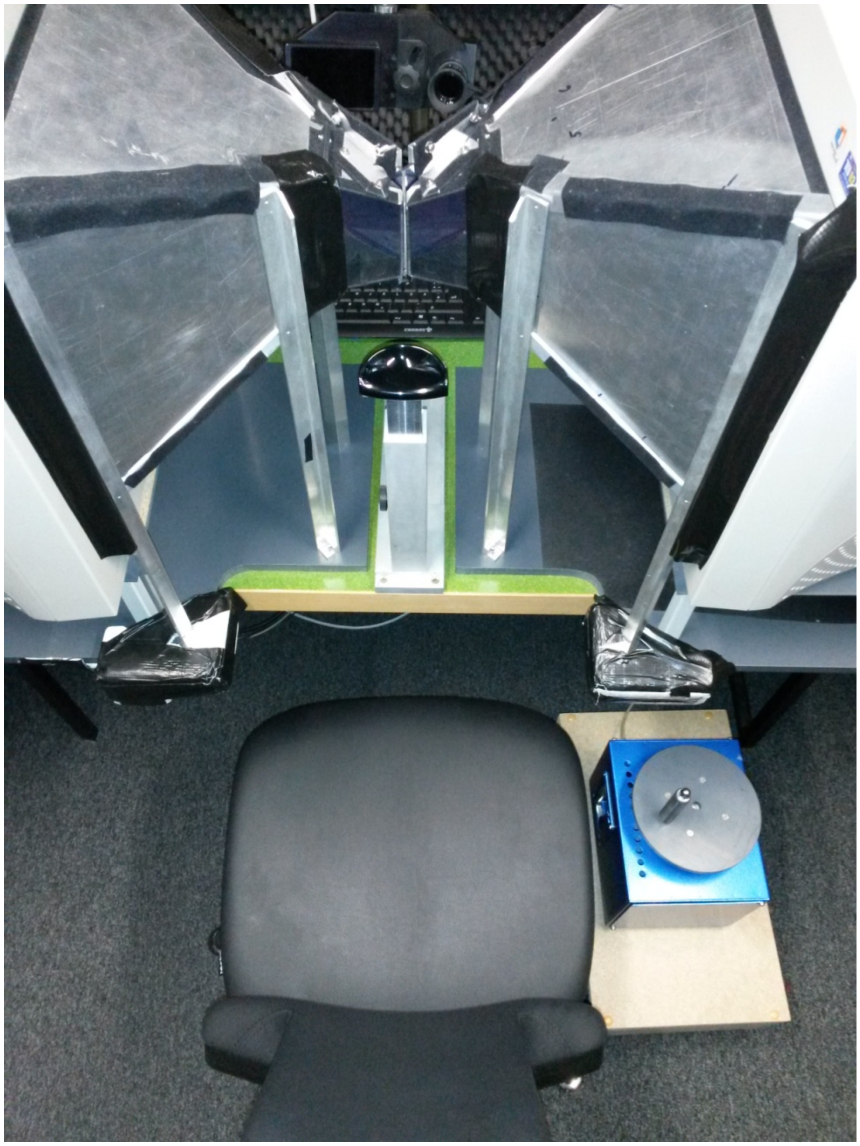

Figure 1. Experimental setup. The cold mirrors of the stereoscope allowed eye-tracking from a table-mounted camera position as seen in the image. The manipulandum device (bottom right) was not visible to participants during the experiment. The experiment was carried out in a dark room.

Each eye's stimulus was centrally presented within a red annulus (inner diameter: 34 degrees), to ensure fusional vergence. The suppressor stimulus completely filled this ring and consisted of overlapping rectangles of random size and color, presented with a refresh rate of 10 Hz. The target stimulus was a horizontally transposing, red, sine-wave grating on black background. The grating's spatial frequency was 0.18 cycles per degree and its total diameter extended 28 degrees. A Gaussian mask, centered within the red annulus, decreased the grating’s luminance towards the edges (Fig. 2).

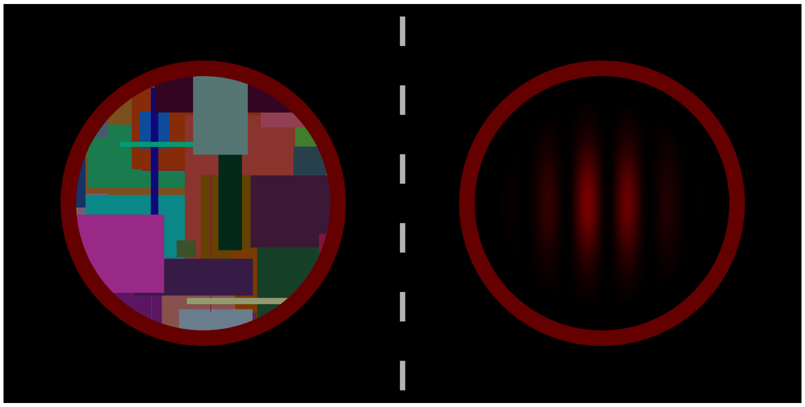

Figure 2. Stimulus. Stimulus display. The dashed line was added to indicate the separation between the two eyes' stimuli.

Procedure

The experiment consisted of three conditions, presented in separate blocks in an alternating fashion. In all conditions, the task of participants was to try and perceive the moving grating stimulus as dominant. They were also instructed to indicate their percepts by keeping a button pressed all the time when full dominance of the target was achieved, and released when the suppressor was dominant. For all participants, button presses were executed with the left hand, while the manipulandum was operated with the right.

In the "coupled action" (CA) condition, participants were instructed to continually rotate the manipulandum in the predefined direction at the velocity of their choosing. Grating velocity was directly coupled to this action at a fixed rate.

In the "decoupled action" (DA) condition, participants executed the same task as in the CA condition, but their action was largely decoupled from the grating's dynamics. The dynamics of the last completed CA block in the experiment was replayed and averaged with the participant's concurrent action with a weight of 4:5. By the concurrent action having only a weight of 1:5, no moment-by-moment prediction could be made regarding the velocity of the grating, while the participant still had some effect on the overall dynamics of the stimulus.

In action blocks (conditions CA & DA), continuous rotation in the correct direction was invoked. If the participant stopped rotating or rotated in the wrong direction, a red rectangle in the center of the stimuli on both screens warned them. Block #2 (first DA block) of participant #12 was removed from analysis due to a failure of following the rotation instructions for 92.04 % of the total duration of the block. In all other blocks and participants, the ratio of erroneous rotation remained negligible (mean = 0.55 %, *SD* = 0.76 %).

In the "no action" (NA) condition, participants were instructed to perform no rotation, while keeping their hands on the manipulandum lever. Stimulus dynamics were entirely determined by the replay of the last completed CA block.

Experimenter and participant only knew whether action was required or a NA block was conducted, but were both blinded to whether a block, in which manual action was required, was CA or DA. Participants in addition were naïve to the fact that two different blocks with movement existed.

The grating's direction was counterbalanced between blocks, while the directional pairing between grating translation and manual action was locked: clockwise action was coupled with leftwards and counter-clockwise action with rightwards motion of the grating stimulus, such that the grating corresponded to the movement when the lever was in front.

Each participant started the experiment with a training block, where the grating stimulus was first introduced separately. This assured that all participants had the same amount of knowledge as for what the target looked like, regardless of their subsequent performance in the binocular task. Following the training, they completed twelve blocks, with block order following a counterbalanced design between the three conditions (four blocks each). Each block was preceded by an eye tracker calibration and lasted 200 s. In those cases when the experiment started with DA or NA blocks (thus no action of the given participant was recorded yet), stimulus dynamics of the last CA block of the previous participant were replayed. For all other DA and NA blocks, recordings of the last CA block of the given participant were used. This allowed us to have a counterbalanced design while maximizing the similarity in stimulus dynamics between conditions.

Analyses

Button press responses were analyzed by comparing conditions in the following measures: predominance of the grating (overall dominance rate throughout the whole experiment), dominance durations (average length of individual dominance periods) for the grating stimulus and for the suppressor, and switching frequency between the two stimuli. Eye data were analyzed by similarly comparing the mean gain of the OKN slow phase (eye velocity as a portion of stimulus velocity). This measure was achieved by, first, removing OKN fast phases and blinks, using the in-built Eyelink saccade detection software with the parameters of 30 degrees/s velocity threshold and 8000 degrees/s^2

acceleration threshold. The average horizontal velocity of OKN slow phases was then calculated by least-squares fitting a linear function to each of the remaining sections. Giving this value the appropriate sign (positive if the slow phase was in the direction of the stimulus in the given block) and dividing it by stimulus velocity yielded the gain of eye movements, at each time point of the experiment where no saccade or blink occurred (Fig. 3). Note that the gain is negative, if the slow phase of the OKN is directed opposite to the grating's drift direction, and positive if both are in the same direction. Perfect OKN would imply a gain of 100 %.

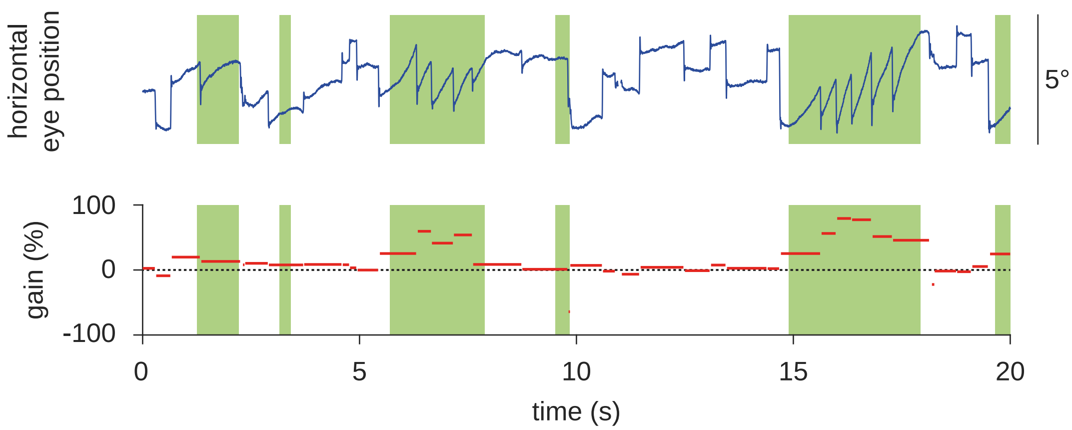

Figure 3. Example excerpt of participant #6. Button press (green areas across both graphs; indicating subjective perceptual dominance of the grating), raw data of horizontal eye movements (blue line), and gain (red line; least-squares mean velocity of slow phase eye movements as portion of stimulus velocity).

Measures were averaged across blocks for each condition and participant, and compared in a within-subjects repeated measures ANOVA. In cases where Mauchly's test indicated a violation of the sphericity assumption, Greenhouse-Geisser corrected values were used. Post-hoc *t* tests were carried out between all condition-pairs, if the variance analysis showed a main effect at a 0.05 alpha level. For post-hoc tests, significance was asserted only when the *p* value fell below a Bonferroni-adjusted alpha-level of 0.05/3 = 0.0167. For analyses on button press measures, blocks with no button press (12 blocks altogether across all participants) were treated as missing values.

Time-normalized analysis

For a more detailed view of temporal relationships between eye movements and the reported percept, we compared the gain for each condition on a scale of normalized time between button press and release. The absolute durations of individual segments from one perceptual switch to the other are variable between participants as well as between the two perceptual interpretations (dominance of grating or suppressor). In order to compare conditions and perceptual dominance without a bias by the absolute duration of the percept, gain data for individual intervals between perceptual switches were normalized to unit time (see also Einhäuser, Martin & König, 2004). The gain trace between each button release and subsequent button press (and vice versa) was mapped by linear interpolation to 10000 data-points referred to as interval [0, 1], prior to averaging these segments. In this time frame, unit time corresponds to one half cycle between two consecutive button presses.

To circumvent the multiple comparisons problem in studying a large number of points on the time-scale, conditions were analyzed in a between-trials cluster-based non-parametric test on the time-normalized data, separately for intervals where the grating or the suppressor was dominant. Methods were based on Maris & Oostenveld (2007). In short, t values were obtained from a pair-wise comparison of conditions at all data points of the time normalized scale. Clusters were formed where paired t tests resulted in statistics with $t > 2$. Gaps between clusters were ignored when they did not exceed 0.5 % of the normalized time scale (corresponding to an average of 44.2 ms in the real-time data). Note that these thresholds do not affect the false alarm rate in the non-parametric test, they only set the sensitivity for localizing the clusters. The sum of t values in the largest cluster was recorded as the observed test statistic. To form random partitions, trials (averages across blocks per participant) of the different conditions were collected in a single data set, and new subsets of equal sizes were formed by random draws of trials. Test statistics were calculated on these random subsets similarly to that on the actual conditions. By repeating the above method of random partitions 1000 times, Monte Carlo p values were calculated by taking the proportion of random partitions that showed a larger test statistic than the observed test statistic.

Results

Subjective report and gain

Button press responses showed no significant difference between conditions in any of the examined variables. Gain evinced to be different between conditions when data were considered irrespective of button press data (Table 1). Here, pair-wise comparisons revealed a significant difference between CA and NA conditions ($t(23) = 2.79$, $p = 0.010$) but not between CA and DA ($t(23) = 1.72$, $p = 0.099$) or DA and NA ($t(23) = 1.41$, $p = 0.173$) conditions. In post-hoc contrasts, a linear model showed the best fit to the pattern of results (Fig. 4; $F(1, 23) = 7.78$, $\eta_p^2 = 0.25$, $p = 0.010$).

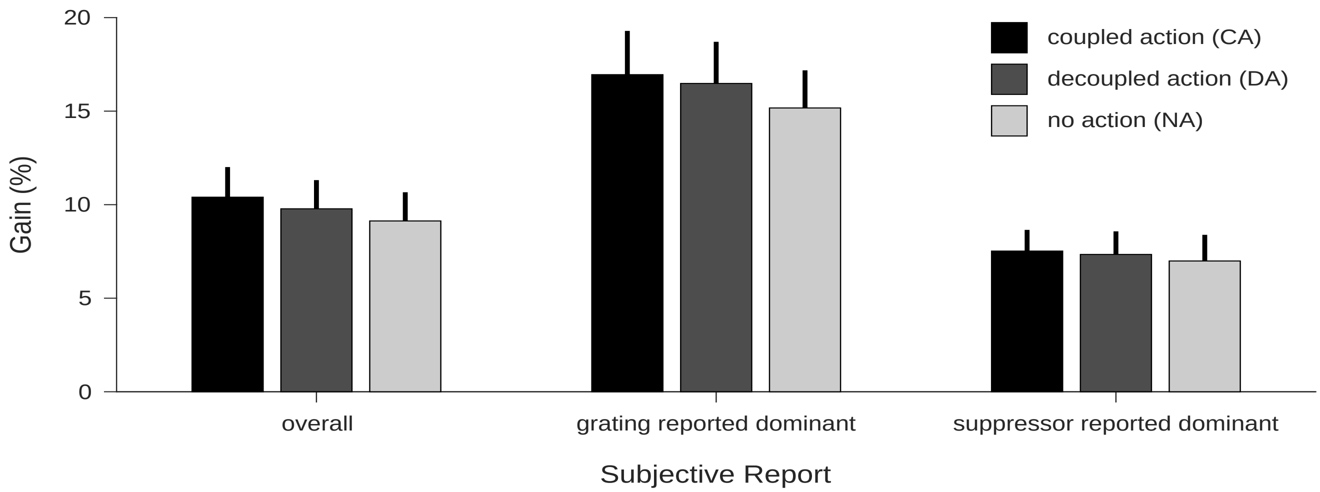

Figure 4. Means of gain. Groups from left to right: data across the whole experiment, portion of experiment when button press report indicated dominance of the grating, and portion of experiment when button press report indicated dominance of the suppressor. Error bars show standard errors of the mean.

Variable	Unit	Condition CA	DA	NA	*df*	*F*	η_p^2	*p*
Predominance (grating)	%	38.47 (25.93)	38.82 (28.80)	38.05 (28.56)	2, 46	0.17	0.007	0.844
Dominance duration (grating)	s	7.74 (11.81)	8.64 (14.05)	8.73 (14.61)	1.32, 30.25	0.39	0.017	0.595
Dominance duration (suppressor)	s	8.61 (6.45)	8.49 (8.28)	10.04 (9.97)	2, 46	0.73	0.031	0.486
Switching frequency	1/s	0.21 (0.15)	0.20 (0.15)	0.21 (0.17)	1.40, 32.24	0.69	0.029	0.459
Gain (overall)	%	10.40 (7.89)	9.78 (7.51)	9.13 (7.52)	2, 46	4.40	0.161	0.018*
Gain (grating dominant)	%	16.95 (11.46)	16.48 (10.91)	15.17 (9.86)	1.54, 35.49	2.84	0.110	0.084
Gain (suppressor dominant)	%	7.52 (5.55)	7.34 (6.05)	6.99 (6.85)	2, 46	0.86	0.036	0.432

Table 1. Statistical measures of results from perceptual reports and eye movements (gain). Means and standard deviations are shown for all variables and conditions (CA: coupled action, DA: decoupled action, NA: no action). Predominance of grating: portion of the experiment's total duration, when perceptual dominance of the grating was indicated by button press. Dominance duration: average time of dominance between reported perceptual switches. Switching frequency: number of switches per second. Gain: slow phase optokinetic nystagmus velocity as portion of stimulus velocity.

Time-normalized analysis

For periods in which participants reported dominance of the grating, non-parametric test results revealed significant differences at a critical alpha-level of $p < 0.05$ between each condition pair ($p_{CA\text{-}DA} = 0.041$; $p_{DA\text{-}NA} = 0.034$; $p_{CA\text{-}NA} = 0.016$). In contrast, for periods in which participants reported the flash to be dominant, only the CA-NA comparison reached significance ($p_{CA\text{-}DA} = 0.103$; $p_{DA\text{-}NA} = 0.120$; $p_{CA\text{-}NA} = 0.048$). More interestingly, the largest clusters of difference occurred at dissimilar temporal positions relative to the reported dominance switches (Fig. 5). For comparisons between action conditions (CA and DA) and the NA conditions, clusters were found where the gain was highest (proximal to when the grating achieved dominance). However, CA and DA conditions differed the most during times when gain was low (nearby the end of the grating's dominance).

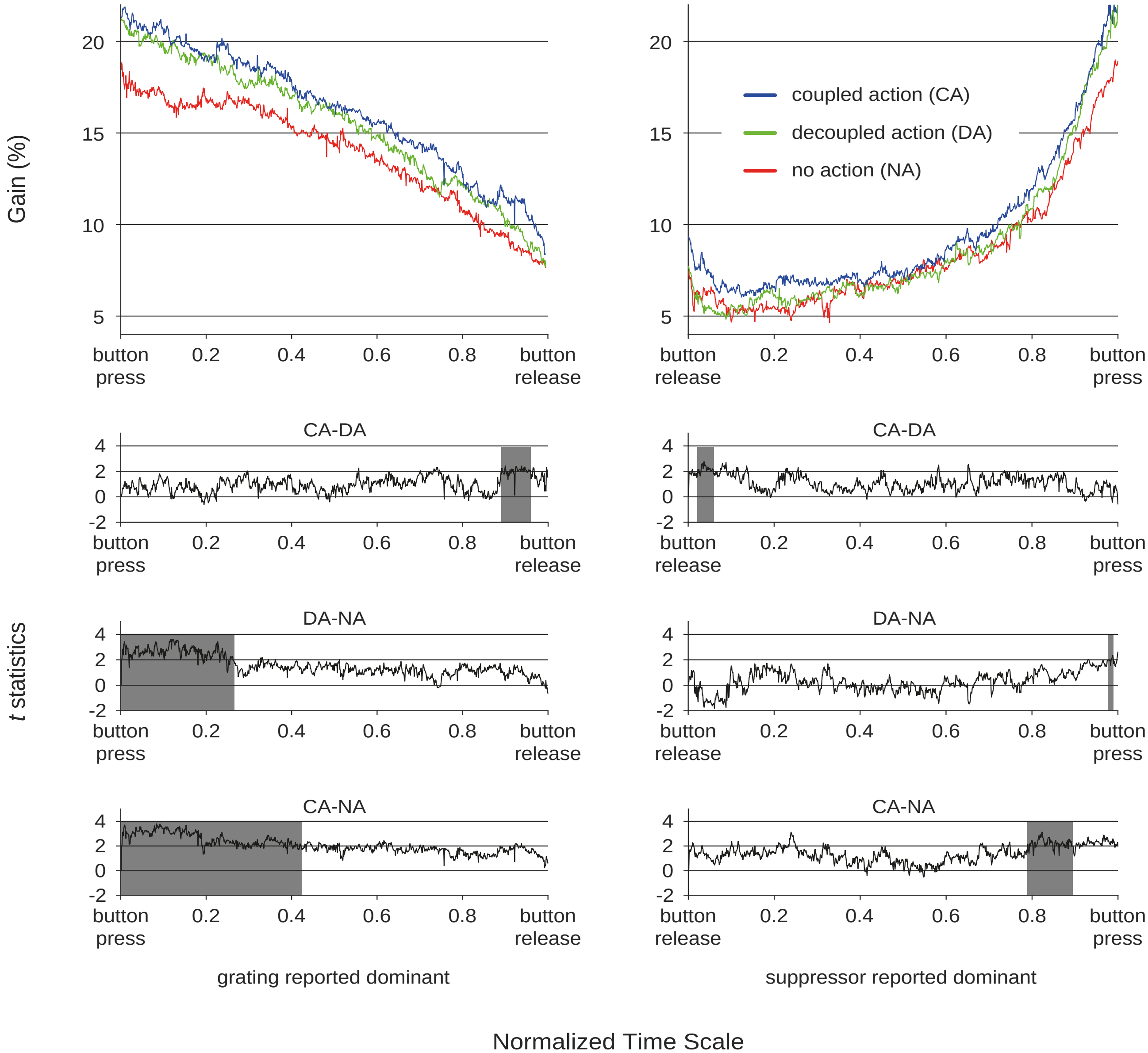

Figure 5. Means of gain, *t* values, and largest clusters on a normalized time scale. *Left panels:* intervals during which button presses indicated dominance of the grating, horizontal (time) axes are normalized such that this interval is mapped to unit time, with 0 being the time of press and 1 the time of release. *Right panels:* intervals during which button presses indicated dominance of the suppressor, time axes normalized to unit time from button release (t=0) to subsequent press (t=1). Means of gain are plotted based on individual segments of dominance normalized to this time scale. Conditions are compared in paired samples *t* tests across time. Test statistics are plotted for each condition pair, with the largest clusters of difference highlighted in grey (for details, see *Analyses*).

Discussion

Our results show that reported percept and eye-movement behavior are affected differently by action. Although earlier studies with ambiguous displays (Beets *et al.*, 2010; Mitsumatsu, 2009; Wohlschläger, 2000) and BR (Di Pace & Saracini, 2014; Maruya *et al.*, 2007) have shown a clear effect of concurrent, task-related action on perception, we did not find any sign of such an effect in a CFS paradigm. This is consistent with the top-down nature of action capture (Wohlschläger, 2000), as the effect of selective attention is weaker in CFS compared to BR, and minimal compared to ambiguous figures (Meng & Tong, 2006; Tsuchiya & Koch, 2005; Tsuchiya, Koch, Gilroy & Blake, 2006).

Visually induced eye-movements, on the other hand, were affected by action, pointing to an underlying mechanism that is either different from that behind the action-to-perception transfer onto the conscious percept, or relies on the same mechanism albeit with diverging sensitivities. While comparing the means of the gain during times of dominance of one or the other stimulus did not show clearly whether the effect stemmed from intervals of the experiment when dominance of the grating was indicated or from the portion when the suppressor was dominant, non-parametric test results showed that the effect was not distributed homogeneously across the time of dominance of either stimulus. As opposed to binary button press responses, the gain of OKN eye movements showed a gradual transition of dominance between suppressor and target stimulus (cf. Fahle, Stemmler, & Spang 2011; Naber *et al.*, 2011). Relative to these transitions, the effect of (both coupled and decoupled) action as compared to no action was most expressed when eye movements elicited by the grating were highest. However, coupled and decoupled action conditions differed from one another mostly at times when the gain of OKN was at its lowest (Fig. 5). This means that not only were eye movements affected by action independently of the reported percept of the stimulus, but also that most of the effect of coupling between action and stimulus dynamics took place when the coupled stimulus was least likely to be in awareness.

Our paradigm tackles an earlier stage of competition than all previously reported experiments on the action-to-perception transfer. As such, the measures that we used reflect a stronger influence of bottom-up processes, while selective attention could less readily bias the percept. This is feasible to be reflected in the lack of an on-line action to perception effect on the reported percept in our results. Eye movements, on the other hand, respond to a lower level of shared processing between action and perception, and display a gradual effect of action – the degree of coupling corresponds to the extent of OKN response to the grating‘s motion.

Although eye movements are a type of motor behavior, in our experiment they are inseparable from the visual information that originates from the target stimulus. It remains an open question, to what extent the underlying perceptual processes also differ between the tested conditions. Nevertheless, as the OKN response is reliably connected to the percept in BR (Naber *et al.*, 2011), we can assume that lower level (pre-attentive) perceptual processes are affected by action, similarly to eye movements. The difference between our results and those of Maruya *et al.* (2007) and Di Pace & Saracini (2014), on the other hand, further support the notion that BR and CFS are principally different in how they affect selective attention and awareness (Tsuchiya & Koch, 2005; Tsuchiya *et al.*, 2006).

One of the earliest and theoretically most important questions of the study of how action influences perception is whether any observed effect is a result of a direct, low level, information flow from the motor to the perceptual domain, or if higher level representations or even endogenous attention serve as vehicle for the transfer. A collection of results from previous research together indicate that neither side can be clearly excluded from an overarching explanation. Our study takes this approach one step further. On one hand, it shows that in CFS, a paradigm where selective attention plays less role than in ambiguous figures or BR, effects of action on perception can still be observed. On the other hand, it also shows that methods most often used in similar studies, which rely on observers' subjective reports, do not capture the effect. This is an indication that conscious perception is not affected at the measured level of perceptual organization.

Acknowledgements

This project was supported by the German Research Foundation, International Research Training Group IRTG 1901, "The Brain in Action". We also thank Adrien Pfeuffer, Alexander Platzner and Matthias Born for their technical support with the manipulandum, and Monique Michl for running the experiment. We thank Marius 't Hart, Denise Henriques and Liad Mudrik for helpful discussion.

References

Beets, I. A. M., 't Hart, B. M., Rösler, F., Henriques, D. Y. P., Einhäuser, W., & Fiehler, K. (2010). Online action-to-perception transfer: Only percept-dependent action affects perception. *Vision Research, 50*, 2633-2641.

Boring, E. G. (1930). A new ambiguous figure. *American Journal of Psychology, 42*, 444-445.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433-436, doi:10.1163/156856897X00357

Brascamp, J. W., Klink, P. C., & Levelt, W. J. (2015). The 'laws' of binocular rivalry: 50 years of Levelt's propositions. *Vision Research, 109,* 20-37.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink toolbox: Eye tracking with Matlab and the psychophysics toolbox. *Behavior Research Methods, Instruments, & Computers, 34,* 613-617, doi:10.3758/BF03195489.

Di Pace, E., & Saracini, C. (2014). Action imitation changes perceptual alternations in binocular rivalry. *PloS ONE, 9*(5), e98305

Einhäuser, W., Martin, K. A. C., & König, P. (2004). Are switches in perception of the Necker cube related to eye-position? *European Journal of Neuroscience, 20*(10), 2811-2818.

Fahle, M. W., Stemmler, T., & Spang, K. M. (2011). How much of the "unconscious" is just pre-threshold? *Frontiers in Human Neuroscience, 5*(120), 1-6.

Keetels, M., & Stekelenburg, J. J. (2014). Motor-induced visual motion: hand movements driving visual motion perception. *Experimental Brain Research, 232,* 2865-2877.

Levelt, W. J. M. (1965). *On binocular rivalry.* Soesterberg, The Netherlands: Institute for Perception RVO-TNO.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164,* 177-190.

Maruya, K., Yang, E., & Blake, R. (2007). Voluntary action influences visual competition. *Psychological Science, 18*(12), 1090-1098.

Meng, M., & Tong, F. (2006). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision,* 4, (7):2, 539-551, http://journalofvision.org/4/7/2/, doi:10.1167/4.7.2.

Metzger, W. (1934). Beobachtung über phänomenale Identität. *Psychologische Forschung, 19,* 1-60.

Mitsumatsu, H. (2009). Voluntary action affects perception of bistable motion display. *Perception, 38*, 1522-1535.

Müsseler, J. (1999). How independent from action is perception? An event-coding account for more equally-ranked crosstalks. In G. Ascherleben, T. Bachman, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 121-147). Amsterdam: Elsevier.

Naber M., Frässle S., & Einhäuser W. (2011). Perceptual rivalry: Reflexes reveal the gradual nature of visual awareness. *PLoS One*, 6, e20910.

Necker, L. A. (1832). Observations on some remarkable optical phaenomena seen in Switzerland; and on an optical phaenomenon which occurs on viewing a figure of a crystal or geometrical solid. London Edinburgh Philosophical Magazine and Journal of Science, 1(5), 329–337.

Pelli, D. G. (1997). The video toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437-442, doi:10.1163/15685689X00366.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology, 9,* 129-154.

Schütz-Bosbach, S., & Prinz, W. (2007). Perceptual resonance: Action-induced modulation of perception. *Trends in Cognitive Sciences, 11*(8), 349-355.

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience, 8*(8), 1096-1101.

Tsuchiya, N., Koch, C., Gilroy, L. A., & Blake, R. (2006). Depth of interocular suppression associated with continuous flash suppression, flash suppression, and binocular rivalry. *Journal of Vision,* 6, (10):6, 1068-1078, doi:10.1167/6.10.6.

Wexler, M., Kosslyn, S. M., & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition, 68,* 77-94.

Wexler, M., & van Boxtel, J. J. (2005). Depth perception by the active observer. *Trends in Cognitive Sciences, 9*(9), 432-438.

Wheatstone, C. (1838). Contributions to the Physiology of Vision. Part the First. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London, 128*, 371-394.

Wohlschläger, A. (2000). Visual motion priming by invisible actions. *Vision Research, 40,* 925-930.

Wohlschläger, A., & Wohlschläger, A. (1998). Mental and manual rotation. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 397-412.

Zwickel, J., & Prinz, W. (2012). Assimilation and contrast: the two sides of specific interference between action and perception. *Psychological Research, 76*(2), 171-182.

Study III.

What you see is what you expect: Cognitive assumptions influence the action-to-perception transfer in ambiguous perception

(submitted manuscript)

Peter Veto[1,2], Marvin Uhlig[1], Nikolaus F. Troje[3], Wolfgang Einhäuser[1]

[1]Chemnitz University of Technology, Institute of Physics - Physics of Cognition Group, Chemnitz, Germany

[2]Philipps-University, Marburg, Germany

[3]Department of Psychology, Queen's University, ON, Canada

Abstract

Can cognition penetrate action-to-perception transfer? Participants observed a cylinder of ambiguous rotation direction. Beforehand, they experienced one of two mechanical models: an unambiguous cylinder was connected to a rod by either a belt (cylinder and rod rotating in the same direction) or gears (both rotating opposite). During ambiguous cylinder presentation, mechanics and rod were invisible. Observers inferred the rod's direction from their moment-by-moment subjective perceptual interpretation of the cylinder. Observers reported the (hidden) rod's direction by rotating a manipulandum in either the same or the opposite direction. For the "belt" model, same-direction report induces congruency between cylinder perception and manual action. Here, same-direction movement stabilized the perceived direction, replicating a known congruency effect. For the "gear" model, opposite-direction report induces congruency between perception and action. Here, no congruency effect was found: perceptual congruency and cognitive model nullified each other. Hence, observers' internal models of a machine's operation guide action-to-perception transfer.

Keywords: vision, action, ambiguous perception, action-perception coupling, cognitive representation, action-to-perception transfer

Introduction

Actions typically have perceptual consequences. Pushing a mouse forward makes a pointer go up, pulling a control stick backward makes a plane climb, pulling a cord down opens the blinds, turning a key releases a lock, turning a screw pushes it forward or backward, etc. These mappings from action to outcome are vastly different, sometimes even conflicting, yet, they appear nearly self-evident to us. Apparently, we have internalized models of the complex mappings between actions and their effects and can recruit them in a context-specific manner. In the present study, we ask whether these cognitive models of action consequences penetrate into perception itself.

Besides affecting perception through changing the external world, action may also directly impact internal perceptual representations. Practicing a movement leads to improved visual discrimination of the same movement[1], action and the perception of action may rely on the same primitives[2], and as humans, we are all "experts" on biological motion perception[3,4]. The notion of shared action-perception representations has been formalized as the theory of common coding[5] and extended into the theory of event coding[6].

Ambiguous stimuli present an excellent means of isolating *direct* effects of action on perception from effects that are mediated through changes in the outside world. For example, when two identical disks move across the screen on the same trajectory but in the opposite direction[7], the direction of a concurrently performed hand action biased the percept to either the two disks moving across or bouncing off one another[8]. Wohlschläger[9] demonstrated that planning or executing a hand movement biased a rotating ambiguous motion display in the direction of manual rotation. Similarly, the perceptual interpretation of an ambiguous (bi-stable) rotating cylinder was stabilized, when viewers reported their perceived direction with congruent manual rotation[10]. Comparable results were found when instead of ambiguous displays, binocular rivalry[11] or unambiguous stimuli with high perceptual uncertainty[12] were used. These studies describe a congruency effect, whereby a match between action and perception (e.g., rotation in the same direction) leads to increased perceptual stability, as compared to an incongruent relationship (rotation in the opposite direction).

Although the theory of common coding[5] can accommodate all of the aforementioned findings, it does not strictly specify the nature of the information transfer from action to perception. One possibility is that the action-to-perception transfer is mediated by cognition. Maruya, Yang and Blake[11] found effects of action on perception, when the action-coupled stimulus was perceptually suppressed (i.e., outside of awareness), arguing against a purely cognitive effect. In contrast, Beets and colleagues[10] found effects only when the action was used to report the perception, suggesting the need for some cognitive component.

In the present study, we used an ambiguous motion stimulus to assess the impact of cognition on action-to-perception transfer. Specifically, we tested whether an observer's internal model of the particular mechanics that link the action to the observed visual consequence modulates the effect of action on perception. In all experimental conditions, observers viewed an ambiguous rotating cylinder (figure 1b; Beets *et al.*[10]) and reported perceived spinning direction by rotating a manipulandum lever either in the same or in the opposite direction as the lever of the assumed mechanism would move (figure 1c, d).

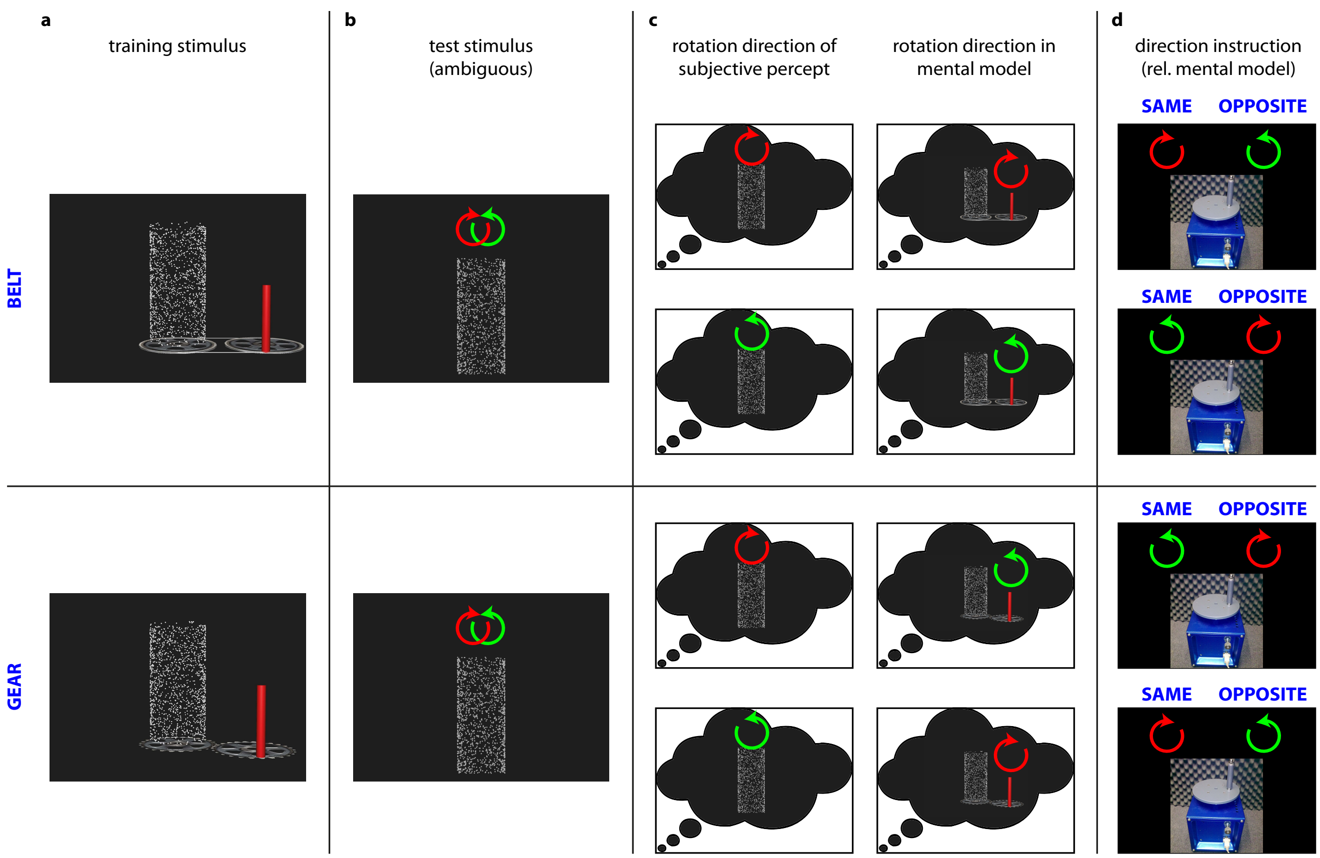

Figure 1. Stimuli, percept, and task. a) First, in separate blocks (blocks 1 & 6, see table 2), participants were introduced to the mechanical model ("belt" or "gear" layout). For 30 seconds, they controlled the displayed motion with the manipulandum. Then, 20 seconds of unambiguous motion followed (the cylinder and mechanical model rotated with occasional switches in direction), where

observers had to report the rotation of the red lever (the red rod attached to the wheel) in accordance with the subsequent experimental block ("same direction instruction" or "opposite direction instruction"). For the last 20 seconds of training, the red lever disappeared and the mechanics was covered by a virtual desk, while the task remained unchanged. **b)** All test blocks showed the same, ambiguous, motion cylinder for 3 minutes each. **c)** Two possible perceptual interpretations of the test stimulus (clockwise and counter-clockwise). Participants had to respond to the imagined motion of the red lever, as it related to their current percept. **d)** Instruction (manipulandum rotation in the same or opposite direction as that of the red lever in the mental model). Note that in the "belt" condition, the same/opposite direction instruction leads to congruency/incongruency between perceived and performed rotation, while this relationship is reversed in the "gear" condition.

Ambiguous stimuli evoke a percept that switches back and forth between perceptual interpretations (here: one direction of rotation or the other). A more stable percept translates to longer periods between these switches. Hence perceptual stability is operationalized as the median duration for which a percept (of either rotation direction) was perceived.

Studies like Wohlschläger[9] or Beets and colleagues[10], by design, confound two effects: (i) the effect of coupling between the perceived direction of the ambiguous motion and the manipulandum rotation and (ii) the effect of coupling between the internal model of the rotation and the executed movement. There is an implicit assumption that the cognitive model and the perception are closely matched. Here we separate these two effects: as in Beets and collagues[10], we used an ambiguous cylinder (figure 1b) and asked observers to rotate a manipulandum in the same or in a different direction as an ambiguously rotating object. However, observers did not report the perceived direction of the ambiguous cylinder itself, but of a visual representation of the manipulandum lever. Observers were taught that the ambiguous cylinder and the lever were either coupled through a belt or through gears (figure 1a). This results in four (2x2) conditions (table 1): the *internal model* (levels: "gear", "belt") and the *match* between perceived cylinder rotation and manipulandum rotation (levels: "congruent", "incongruent").

During testing, the visual representation of the lever is not shown (figure 1b), rendering all four conditions visually identical. If the congruency between perceived motion direction and manipulandum rotation is decisive for action-perception coupling, the effect of the *match* between action and perception on perceptual stability should be independent of the *internal model* (figure 2, *left*). If, however, the cognitive congruency between the invisible representation of the lever and actual lever is decisive, we expect the effects to reverse for the gear model (figure 2, *right*). Even if the effect does not reverse completely, any interaction between *internal model* and perception-action *match* would point to cognitive penetration of action-perception coupling.

internal model	*match* (perceived rotation – manual action)	instruction	effect if perception dominates	effect if internal model dominates
belt	congruent	same dir.	stabilize	stabilize
belt	incongruent	opposite dir.	destabilize	destabilize
gear	incongruent	same dir.	destabilize	stabilize
gear	congruent	opposite dir.	stabilize	destabilize

Table 1. Conditions. There are 4 (2x2) experimental conditions, defined by the factors *internal model* ("belt", "gear") and *match* ("congruent", "incongruent"). Note that *match* and instruction in the belt and gear conditions are inversely related. Depending on whether perceptual congruency or internal model dominates, different predictions on perceptual (de)stabilization result (right columns).

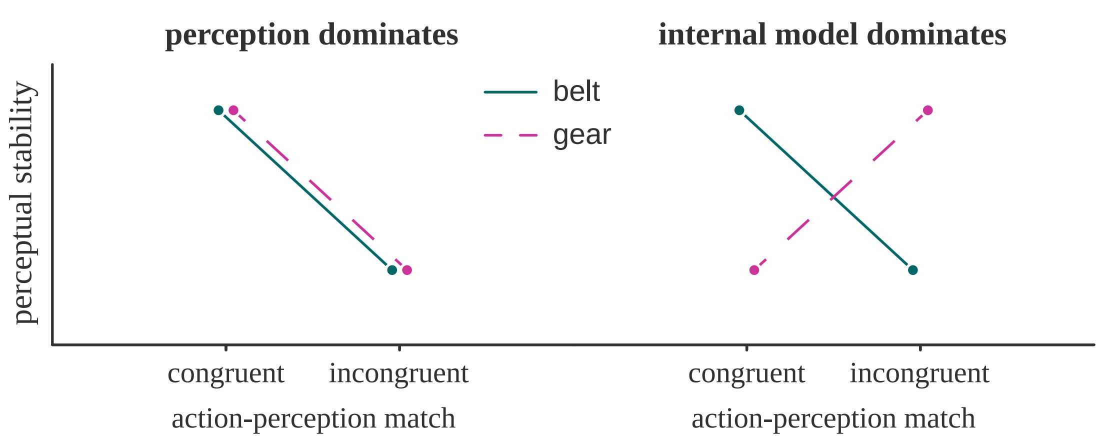

Figure 2. Hypotheses. Expected effects of *internal model* and *match* (between percept and action) on perceptual stability, if action-perception coupling is not under cognitive influence (*left*) or dominated by the cognitive model (*right*).

Results

We measured perceptual stability and tested whether the factors *internal model* [gear vs. belt] and *match* [congruent vs. incongruent] influenced this measure. We found no significant main effect of either factor (*internal model*: $F(1, 31) = 0.517$, $\eta_p^2 = 0.016$, $p = 0.478$; *match*: $F(1, 31) = 2.697$, $\eta_p^2 = 0.080$, $p = 0.111$), while the two factors showed a significant interaction ($F(1, 31) = 4.763$, $\eta_p^2 = 0.133$, $p = 0.037$). Paired samples two-sided *t*-tests revealed that the congruency effect was significant only in the "belt" condition (perceptual stability: $\text{mean}_{\text{Belt Congruent}} = 4.40$ s; $\text{mean}_{\text{Belt Incongruent}} = 3.24$ s; $t(31) = 2.759$, $p = 0.010$), but not in the "gear" condition ($\text{mean}_{\text{Gear Congruent}} = 3.39$ s; $\text{mean}_{\text{Gear Incongruent}} = 3.67$ s; $t(31) = 0.661$, $p = 0.513$; figure 3).

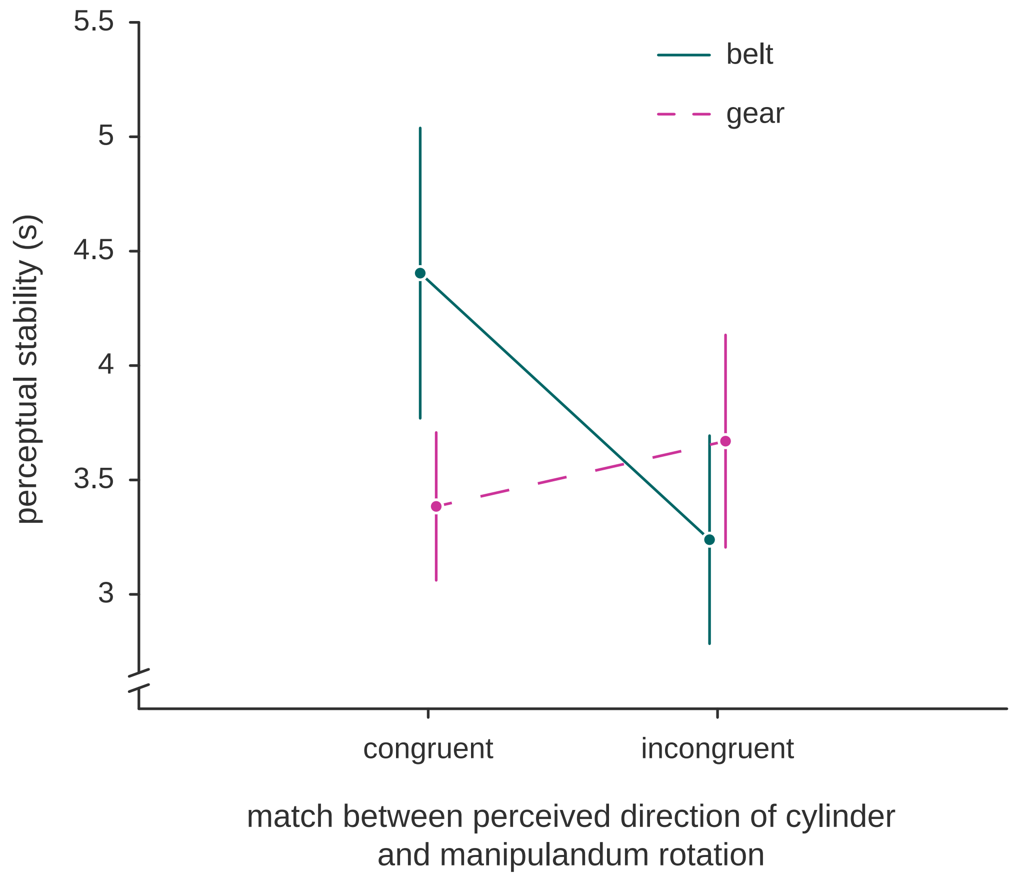

Figure 3. Results. Perceptual stability for each condition, averaged across participants (N=32; mean perceptual stability across observers, where perceptual stability for each observer is the median duration for which a percept was perceived). Error bars show standard errors of the mean.

Discussion

The interaction between the factors *internal model* and *match* revealed that the cognitive model of the coupling between an action and its observable effect significantly influenced the action-to-perception transfer. Results from the "belt" condition replicated the known congruency effect[10], while the lack of effect in the "gear" condition (in the presence of an interaction) showed that the influence of the assumed mechanical model counteracted the natural congruency bias. Thus, cognition plays a significant role in action-to-perception transfer, while it is not the sole source of the effect.

In the framework of the common-coding theory[5], our results can be interpreted as evidence that the shared representations between perception and action occur on a cognitively accessible level of processing. This is in line with the observation that action-perception transfer can depend on the relevance of an action for the perceptual task[10]. A direct influence of the cognitive model on action-to-perception transfer might also be of adaptive value in real-life situations, in particular when tools similar to the one used here are involved: evoking a cognitive model allows better predictions of an action's consequences and may therefore result in better performance or quicker learning of a complex manual task[13]. Nonetheless, our results do not exclude that on some level shared action-perception representations exist that are under less cognitive control and form independently of awareness[11]. This is particularly conceivable for low-level representations, where perception itself may remain unaffected by either action or the cognitive expectation of perception-action coupling[14]. In a representation-based framework, the results on action-perception transfer taken together necessitate different representational levels, of which only some are penetrable to executive functions, awareness or cognition.

A complementary view posits that the quality of perception arises from the relation between our actions and their sensory effects[15]. Perception, cognition and action then become intimately related through the model that is generated by observing the sensory consequences of an action. In this case, perceptual qualities *and* the cognitive model can be viewed as consequences of the action-perception relation. This is consistent with a recently proposed action-oriented framework, in which perception and cognition are formed together, with action being the key organizing force behind both[16]. In a simple system like the gear/belt mechanics, it would appear conceivable that instruction led the observers to simply learn the coupling from action-to-perception without forming a cognitive model. However, we deliberately chose an experimental design that reversed the congruency instruction without re-exposing the observers to the action-perception coupling; instead, we exposed each observer to their second mechanical model only after he or she had completed all blocks with the first model (order of mechanics balanced across observers; table 2). Observers therefore needed to apply their internal mechanical model to reverse the instruction without practicing the action-perception contingency. This makes it likely that observers indeed have formed a cognitive model during instruction, which they consistently applied until a different model became evident through a new, externally available, action-perception contingency.

Methods

Participants

Thirty-two naïve participants (15 males, 17 females; 25 +/- 4.6 years; 4 left- and 28 right-handed) were included in the analysis. In one additional observer, no reversal of their rotation occurred in several blocks; this was detected when visually inspecting data quality, and the data of this observer was excluded prior to any further analysis. One further observer was assigned to a wrong group by technical error, which was realized during the experiment and their data was not analyzed or inspected any further. Procedures conformed to the Declaration of Helsinki and were approved by the Ethikkommission FB04 of Philipps-University Marburg (#2011-04K). Participants gave written informed consent prior to their partaking. All participants had normal or corrected-to-normal vision.

Setup and stimuli

Stimuli were generated using Unity3D (Unity Technologies, San Francisco, CA) and Matlab (Mathworks, Natick, MA) with the Psychophysics Toolbox[17,18], and presented on an Asus GL502 laptop computer (screen resolution: 1920 x 1080 at 60 Hz; viewing distance: 73 cm). Manual responses were recorded by a custom-built manipulandum device (figure 1d shows an image of the device), measuring the angular position of the rotating handle via a Kübler Sendix 5020 incremental rotary encoder. For all participants, the manipulandum was placed on the right side of the chair.

For training blocks, stimuli were rendered with a perspective camera and other depth-cues present. The 3-dimensional model of the cylinder consisted of small spheres placed at equal distances from a vertical axis, with randomly defined vertical and angular positions relative to the axis. The total size of the display extended 14.4 x 11.7 degrees of visual angle, with the diameter of each dot being 0.08 degrees. The mechanical model consisted of either two wheels connected by a belt, or two adjacent cogwheels. The wheels moved according to the type of connection, that is, same direction in the “belt” condition and opposite direction in the “gear” condition. One of the wheels was placed directly below the cylinder and always moved together with it, as if they were fastened together. This wheel, as well as the cylinder, was shown in the center of the screen. The second wheel was to the right, with a vertical red rod attached to the top side (resembling the handle of the manipulandum; see figure 1a).

For test blocks, the same cylinder object was depicted in the center (6.1 x 12.7 degrees) as an orthographic projection, without depth-cues or the attached mechanical model (figure 1b). This way, the direction of rotation was completely ambiguous and up to the perceptual interpretation of the viewer. The spheres of the cylinder were shown in a homogeneous color (appearing as 2-dimensional dots) and their size did not change along their movement trajectory. Thus, the front and rear surfaces of the cylinder were identical and showed no cues of occlusion. However, due to the dynamics of the dot movements, this cylinder formation is consistently perceived as a 3-dimensional rotating object (structure-from-motion), where the apparent direction of rotation is ambiguous and its perception alternates (see e.g. Beets *et al.*[10]).

Procedure

For each participant, the experiment consisted of two halves, one with the “belt” and the other with the “gear” stimulus condition (order counterbalanced between participants). Each half of the whole experiment started with a training block that introduced the stimulus and mechanical model of the applicable condition, followed by four test blocks with the ambiguous stimulus (table 2).

	1	2	3	4	5	6	7	8	9	10
	mechanics training (total: 70s)	**test (180s)**	**test (180s)**	**test (180s)**	**test (180s)**	**mechanics training (total: 70s)**	**test (180s)**	**test (180s)**	**test (180s)**	**test (180s)**
I	belt	S	O	O	S	gear	S	O	O	S
II	gear	S	O	O	S	belt	S	O	O	S
III	belt	S	O	O	S	gear	O	S	S	O
IV	gear	S	O	O	S	belt	O	S	S	O
V	belt	O	S	S	O	gear	S	O	O	S
VI	gear	O	S	S	O	belt	S	O	O	S
VII	belt	O	S	S	O	gear	O	S	S	O
VIII	gear	O	S	S	O	belt	O	S	S	O

Table 2. Design matrix. Within a sequence of test blocks, order of reporting conditions follows an ABBA pattern (either SOOS or OSSO; S: same direction instruction, O: opposite direction instruction). The starting of the sequence and all other variables of the design were counterbalanced between participants, leading to a total of 8 possible block order combinations; that is, each block order (I-VIII) was assigned to four of the 32 observers.

The training blocks were designed to gradually introduce model and task to the participant. In the first 30 seconds, the movement of the unambiguous cylinder and the attached mechanical model were directly connected to the manipulandum lever. Participants were instructed to move the lever as they wished and to observe the mechanical workings of the model. Then, for a 20 seconds interval, the model rotated at a constant velocity, changing direction every 6 +/- 2 seconds. Participants had to either mimic the movement of the red rod on the attached wheel, or rotate in the opposite direction (according to what the instruction would be in the subsequent test block). In the last 20 seconds of the training block, participants continued with their previous task but the red rod disappeared, and the mechanics were occluded by a virtual desk. This way, the movement of the cylinder was still unambiguous, but the task of the participant was already identical to what they would do in the subsequent test block. To make certain that the correct response was practiced, a salient red rod appeared directly to the left of the cylinder, when the response direction was incorrect. Furthermore, the experimenter was also present during the training block and verified that by the end of the instruction, all participants understood the current task.

Test blocks always showed the ambiguous stimulus (figure 1b), moving at a constant velocity (90°/s). Depending on the condition of the given block, participants had to move the manipulandum lever in the same or opposite direction as the red lever on the mechanics (as seen in the training block) would rotate. Test blocks lasted 3 minutes each. Before each block, the starting position of the manipulandum lever was set to the 12 o'clock position.

The order of training stimulus ("belt" or "gear" mechanics) between the two halves of the experiment, the order of test block instructions within one half of the experiment (same or opposite direction; always in an ABBA order), as well as the order of test block instructions between the two halves of the experiment were counterbalanced between participants (table 2).

Analysis

Manipulandum rotation velocity data were segmented into periods of rotation in one direction or the other, as well as periods with no movement (no change of position for at least two subsequent sample points). For each observer and condition, perceptual stability was defined as the median duration of all non-zero velocity segments.

Comparisons between conditions were made using a within-subjects repeated measures ANOVA, with factors *internal model* ("belt" or "gear") and *match* (congruent and incongruent, as in the relation between stimulus and action). Effects were considered significant at a 0.05 alpha level, while a Bonferroni-adjusted alpha-level of 0.05/4 = 0.0125 was used for post-hoc *t*-tests.

All reported tests are two-sided, Matlab (R2015a) was used for data processing, SPSS (version 24) for statistical analysis.

References

1. Casile, A. & Giese, M. A. Nonvisual motor training influences biological motion perception. *Curr. Biol.* **16**, 69-74 (2006).
2. Mataric, M. J. Sensory-motor primitives as a basis for imitation: Linking perception to action and biology to robotics in *Imitation in Animals and Artifacts* (ed. C. Nehaniv & K. Dautenhahn) 391-422 (Cambridge, MA: MIT Press, 2000).
3. Troje, N. F. Biological motion perception in *The Senses: A Comprehensive Reference* (ed. A. I. Basbaum, M. C. Bushnell, D. V. Smith, G. K. Beauchamp, S. J. Firestein, P. Dallos, D. Oertel, R. H. Masland, T. D. Albright, J. H. Kaas & E. P. Gardner) 231-238 (Oxford: Elsevier, 2008).
4. Troje, N. F. What is biological motion?: Definition, stimuli and paradigms in *Social Perception: Detection and Interpretation of Animacy, Agency, and Intention* (ed. M. D. Rutherford & V. A. Kuhlmeier) 13-36 (Cambridge, MA: MIT Press, 2013).
5. Prinz, W. Perception and action planning. *Eur. J. Cogn. Psychol.* **9**, 129-154 (1997).
6. Müsseler, J. How independent from action is perception? An event-coding account for more equally-ranked crosstalks in *Cognitive Contributions to the Perception of Spatial and Temporal Events* (ed. G. Aschersleben, T. Bachman & J. Müsseler) 121-147 (Amsterdam: Elsevier, 1999).
7. Metzger, W. Beobachtung über phänomenale Identität. *Psychol. Res.* **19**, 1-60 (1934).
8. Mitsumatsu, H. Voluntary action affects perception of bistable motion display. *Percept.* **38**, 1522-1535 (2009).
9. Wohlschläger, A. Visual motion priming by invisible actions. *Vis. Res.* **40**, 925-930 (2000).
10. Beets, I. A. M., 't Hart, B. M., Rösler, F., Henriques, D. Y. P., Einhäuser, W. & Fiehler, K. Online action-to-perception transfer: Only percept-dependent action affects perception. *Vis. Res.* **50**, 1633-1641 (2010).
11. Maruya, K., Yang, E. & Blake, R. Voluntary action influences visual competition. *Psychol. Sci.* **18**, 1090-1098 (2007).
12. Keetels, M. & Stekelenburg, J. J. Motor-induced visual motion: Hand movements driving visual motion perception. *Exp. Brain Res.* **232** 2865-2877 (2014).
13. Lupyan, G. Cognitive penetrability of perception in the age of prediction: Predictive systems are penetrable systems. *Rev. Philos. Psychol.* **6**, 547-569 (2015).
14. Pylyshyn, Z. Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behav. Bran Sci.* **22**, 341-423 (1999).
15. O'Regan, J. K. & Noë, A. A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.* **24**, 939-1031 (2001).
16. Engel, A. K., Maye, A., Kurthen, M. & König, P. Where's the action? The pragmatic turn in cognitive science. *Trends Cogn. Sci.* **17**, 202-209 (2013).
17. Brainard, D. H. The psychophysics toolbox. *Spatial Vis.* **10**, 433-436 (1997).
18. Pelli, D. G. The video toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vis.* **10**, 437-442 (1997).

Zusammenfassung

Diese Dissertationsschrift besteht aus drei Studien, die sich mit komplementären Aspekten der Handlungs-Wahrnehmungs-Kopplung beschäftigen. In den letzten Jahrzehnten hat in der kognitiven Psychologie ein Paradigmenwechsel begonnen: an die Stelle des traditionellen Zugangs - zuerst kommt der Reiz, dann folgt die Handlung als bloße Antwort darauf – tritt mehr und mehr ein weniger unidirektionales Bild von Handlung und Wahrnehmung. Handlung beeinflusst Wahrnehmung zunächst auf mehr oder weniger triviale Weise durch ihren Effekt auf die Außenwelt – wir bewegen uns oder unsere Sinnesorgane im Raum oder wir bewegen Objekte. Interessanter ist der interne Einfluss der Handlung auf die Wahrnehmung. Studien II und III dieser Arbeit beschäftigen sich direkt mit diesem Thema, indem sie den Einfluss der Handlung auf die Wahrnehmung physisch unveränderliche Reize untersuchen. Studie I beschäftigt sich mit biologischer Bewegung. In meiner Arbeit lege ich dar, dass die Wahrnehmung biologischer Bewegung einen realitätsnahen Spezialfall direkten Handlungs-Wahrnehmungs-Transfers darstellt. Die Hinweisreize für belebte Fortbewegung werden schnell und aufwandsfrei erkannt und erlauben gleichzeitig eine schnelle Aufnahme detaillierter Information über den Handelnden, da wir hierbei unsere immense Erfahrung mit der Bewegung unseres eigenen Körpers unter Berücksichtigung physikalischer Gesetze zur Interpretation dieser Hinweisreize nutzen können. Zusammengefasst ergeben die Studien dieser Arbeit ein frisches Bild der gemeinsamer Repräsentationen von Wahrnehmung und Handlung, ihrer perzeptuellen Folgen und ihrer Beziehung zu kognitiven Modellen der Welt.

In Studie I zeigten wir, dass biologischer Bewegung die Wahrnehmung der Größe des Handelnden beeinflusst. Ein Reiz, der biologische Bewegung darstellt, wird größer wahrgenommen als ein visuell vergleichbarer Kontrollreiz und lässt nachfolgende Reize kleiner erscheinen. Vor dem Hintergrund der Wichtigkeit biologischer Bewegung ist dieses Ergebnis mit anderen Studien im Einklang, die Wichtigkeit zu wahrgenommener Größe in Beziehung setzen. Die Verbindung zu biologischer Bewegung wurde vor dieser Arbeit noch nicht hergestellt. In Studie II verbanden wir einen Reiz, der sich mit einem anderen in Wettstreit befand, mehr oder weniger stark mit einer gleichzeitig ausgeführten Handlung. Während der Grad der Kopplung zwischen Handlung und Wahrnehmung den Bericht der Versuchsperson über die Sichtbarkeit nicht nachweislich beeinflusste, zeigte sich eine deutliche Modulation okulomotorischer Maße. Dieses Ergebnis legt verschiedene Stufen der Handlungs-Wahrnehmungs-Kopplung auf verschiedenen Repräsentationsstufen nahe, die wiederum unterschiedlichen Zugang zu bewusster Wahrnehmung haben. Beeinflusst umgekehrt das kognitive

Modell der Welt den Grad der Handlungs-Wahrnehmungs-Kopplung? In Studie III zeigten wir, dass der Effekt der Handlungs-Wahrnehmungs-Kongruenz auf die Wahrnehmungsstabilität kritisch vom kognitiven Modell der Handlungs-Wahrnehmungs-Kopplung abhängt. Zusammengenommen zeigen Studien II und III, dass kein einzelner Mechanismus und keine einzelne Repräsentation allein für alle Befunde zur Handlungs-Wahrnehmungs-Kopplung verantwortlich sein können. In der übergreifenden Diskussion werde ich die nötigen Anpassungen existierender Modelle betrachten und alternative theoretische Ansätze aufzeigen.

Acknowledgements

First and foremost, I would like to express my appreciation to Professor Wolfgang Einhäuser-Treyer, who supervised and tirelessly supported me throughout the good and the bad of the past four years. Not only did he show excellent example in carrying out research, but he also has always been very patient with me. Despite the hectic changes during these years, he was always there when I needed help: both in technical questions and in motivating through disappointing events.
I am also very grateful to Professor Nikolaus Troje for having hosted me in the Biomotion Lab in Kingston, supervising two of my projects during the year that I spent there – and even thereafter, when I was already in Chemnitz. His work has been an inspiration since years before I started my PhD and it is a real honor to having been part of his team for a while.
Finally, I would like to thank Professor Anna Schubö and all members of the panel for committing their time and evaluating my work.

I also thank the Deutsche Forschungsgemeinschaft for funding. Being part of the International Research Training Group 1901 “The Brain in Action” has greatly added to my knowledge about scientific topics also outside of my immediate field.

As roommates in my first office as a doctoral candidate, Josef Stoll and Adrien Pfeuffer have helped me tremendously with technological skills, as well as by sharing a lot of wit and comradery. I hope that they are also proud of the fact that the laughter from our room was remarked in the building as exceptional. Besides, my attitudes towards computers have changed probably for ever, due to their positive input.

Last, but by no means least, I would like to thank Dr Paul Hemeren for having started me on the long journey of becoming a researcher. Without him, none of the above would have happened – and I feel truly lucky for the opportunity of being on this path.

(content removed)

(content removed)

(content removed)

(content removed)

(content removed)

(content removed)

Erklärung

Ich versichere, dass ich die vorliegende Dissertation

“The connection between action and perception”

selbstständig und ohne unerlaubte Hilfe angefertigt habe und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.

Diese Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Chemnitz, 23. Januar 2018

Peter Veto